

Result No.	Query No.	Score	Query %		Length	DB	ID	Description
			Match	Match				
1	1	1065	100.0		198	20	AAW98158	Human Bcl-2 intera
2	2	908	85.3		196	20	AAW98156	Murine Bcl-2 inter
3	3	696	65.4		138	20	AAW98164	Human Bim-L mutant
4	4	696	65.4		138	20	AAW98157	Human Bcl-2 intera
5	5	691	64.9		138	20	AAW98105	Human Bim-L mutant
6	6	691	64.9		138	20	AAW98166	Human Bim-L mutant
7	7	691	64.9		138	20	AAW98168	Human Bim-L mutant
8	8	690	64.8		138	20	AAW98167	Human Bim-L mutant
9	9	589	55.3		140	20	AAW98135	Murine Bcl-2 inter
10	10	584	54.8		140	20	AAW98160	Murine Bim-L mutan
11	11	584	54.8		140	20	AAW98161	Murine Bim-L mutan



CC treatment or prophylaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 XX  
 SQ Sequence 196 AA;

Query Match 85.3%; Score 908; DB 20; Length 196;  
 Best Local Similarity 86.5%; Pred. No. 2e-78; 11; Indels 6; Gaps 2;  
 Matches 173; Conservative 10; Mismatches 11; Indels 6; Gaps 2;  
 QY 1 MAKQPSDVSSCEDREGRLQPAERPQLRPGAPTSLSQTEPQGNHGGEGDSCPHGSP 60  
 DB 1 makqpsdvsscedreggqlqaerppqlrpgaptslqtepgnqd----gedrcphgsp 56  
 QY 61 QGLAPPASPGFPATRSPLIFPMRRSLLSRSSSGYFSDTRSPAPMSCDKSTQTPSP 120  
 DB 57 qglappaspgpfatrsplifvrrssllsrsssgyfsfdrspapmscdkstqtpsp 116  
 QY 121 COAFNHLYLSAMASMRQA--EPADMPEIWIQELRRIGDEFNAYYARRVFLNNYQAEDH 178  
 DB 117 cqafnhylsamasirgsqeepeelrpeiriaqelrrigdefnetyrrvfyandyeaeh 176  
 QY 179 PRMVLRLRYIVRLVWRMH 198  
 DB 177 pqmvlrlrlfvlvwrh 196

RESULT 3  
 AAW98164  
 ID AAW98164 standard; Protein; 138 AA.  
 AC AAW98164;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX Human Bim-L mutant D51G.  
 XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; human; cancer; autoimmune disease; therapy;  
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
 KW cytotoxicity; agonist.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Binding-site 42..71  
 FT /note= "dynein light chain binding region, present  
 FT only in Bim-L, Bim-EL"  
 FT  
 FT Misc-difference 51  
 FT /note= "Gly replaces Asp of wild-type Bim-L"  
 FT Region 94..102  
 FT /note= "BH3"  
 FT Region 123..137  
 FT /note= "hydrophobic region"  
 FT  
 XX WO9914321-A1.  
 XX  
 PD 25-MAR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-AU00772.  
 XX  
 XX 24-SEP-1997; 97AU-0009373.  
 PR 17-SEP-1997; 97AU-0009263.

XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PA Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX WPI; 1999-244030/20.  
 DR New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 XX  
 PS Claim 25; Page -; 145pp; English.  
 XX  
 CC The present sequence is a D51G mutant of the long form (L) of  
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also  
 CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is  
 CC capable of inducing cell death (apoptosis) and acts as a  
 CC 'death-ligand' for certain members of the pro-survival Bcl-2  
 CC family. Binding the dynein light chain regulates the pro-apoptotic  
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does  
 CC not bind to dynein light chain, is a much more potent killer than  
 CC either Bim-L or Bim-EL. The invention provides variants (see  
 CC AAW98159-58) of murine and human Bim-L and Bim-EL that cannot bind,  
 CC couple or otherwise associate with a dynein light chain owing to  
 CC amino acid addition, substitution and/or deletion within the region  
 CC which binds the dynein light chain. The variants act as agonists  
 CC of Bim function, interacting with anti-apoptotic molecules such as  
 CC Bcl-2 to prevent their functional activity, thereby promoting  
 CC apoptosis, and can be used e.g. in the treatment of cancer or to  
 CC deplete autoreactive lymphocytes in autoimmune diseases. Since Bim  
 CC is expressed in germ cells, modulating its expression or activity  
 CC may be useful as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the human Bim-L sequence given on page 100-101.  
 XX  
 SQ Sequence 138 AA;

Query Match 65.4%; Score 696; DB 20; Length 138;  
 Best Local Similarity 69.7%; Pred. No. 1.8e-58;  
 Matches 138; Conservative 0; Mismatches 0; Indels 60; Gaps 1;  
 QY 1 MAKQPSDVSSCEDREGRLQPAERPQLRPGAPTSLSQTEPQGNHGGEGDSCPHGSP 60  
 DB 1 makqpsdvsscedreggqlqaerppqlrpgaptslqtepgnqd----gedrcphgsp 56  
 QY 61 QGLAPPASPGFPATRSPLIFPMRRSLLSRSSSGYFSDTRSPAPMSCDKSTQTPSP 120  
 DB 42 -----drspapmscdkstqtpsp 60  
 QY 121 COAFNHLYLSAMASMRQAEPADMPEIWIQELRRIGDEFNAYYARRVFLNNYQAEDH 180  
 DB 61 cqafnhylsamasmrgaepadmpeiriaqelrrigdefnayyarrvflnnyyqaedhpr 120  
 QY 181 MVLRLRYIVRLVWRMH 198  
 DB 121 mvllrllyivrlvwrh 138  
 RESULT 4  
 AAW98157  
 ID AAW98157 standard; Protein; 138 AA.  
 XX AAW98157;  
 AC AAW98157;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX Human Bcl-2 interacting mediator of cell death Bim-L isoform.  
 DE Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; human; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;

isoform.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Binding-site 42..71 /note= "dynein light chain binding region, present  
 FT only in Bim-L, Bim-EL"  
 FT Region 92..100 /note= "BH3"  
 FT Region. 121..135 /note= "hydrophobic region"  
 FT  
 XX WO9914321-A1.  
 XX 25-MAR-1999.  
 XX 17-SEP-1998; 98WO-AU00772.  
 XX 24-SEP-1997; 97AU-0009373.  
 XX 17-SEP-1997; 97AU-0009263.  
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 XX Puthalakath H, Strasser A;  
 XX WPI: 1999-244030/20.  
 XX N-PSDB: AAX24996.  
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 XX treatment  
 XX Claim 15; Page 100-101; 145pp; English.  
 XX The present sequence is the long form (L) of human Bim, or Bcl-2  
 XX interacting mediator of cell death, a novel member of the Bcl-2  
 XX family that is capable of inducing cell death (apoptosis) and which  
 XX acts as a 'death-ligand' for certain members of the pro-survival  
 XX Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology  
 XX region which it encompasses is BH3. It is the only BH3-only protein  
 XX for which splice variants exist. These result in the expression of  
 XX a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL. cDNAs (see  
 XX AAX24996-97) encoding human Bim-L and Bim-EL (see AAW98165) were  
 XX obtained from embryo and liver cDNA libraries using mouse Bim cDNA  
 XX as probe. Murine Bim-S, Bim-L and Bim-EL isoforms are also  
 XX provided (see AAW98154-56). Binding the dynein light chain was shown  
 XX to regulate the pro-apoptotic activity of Bim. Bim-S, the splice  
 XX variant which does not bind to dynein light chain, is a much more  
 XX potent killer than either Bim-L or Bim-EL. The invention provides  
 XX variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that  
 XX cannot bind, couple or otherwise associate with a dynein light  
 XX chain. The identification of Bim permits the identification and  
 XX rational design of a range of products for use in therapy,  
 XX diagnosis, antibody generation and involving modulation of  
 XX physiological cell death. These therapeutic molecules may act  
 XX as either antagonists or agonists of Bim's function and will be  
 XX useful in cancer, autoimmune or degenerative disease therapy.  
 XX Increased Bim expression or Bim activity is useful, e.g. for  
 XX treatment or prophylaxis in conditions such as cancer and deletion  
 XX of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 XX expression of Bim activity is useful in regulating inhibition or  
 XX prevention of cell death or degeneration such as under cytotoxic  
 XX conditions during e.g. gamma-irradiation and chemotherapy or during  
 XX HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 XX hypoxia, degenerative diseases or for prolonging the survival of  
 XX cells being transplanted for treatment of disease. Since Bim is  
 XX expressed in germ cells, modulating Bim expression or Bim activity  
 XX is useful, e.g. as a contraceptive or method of sterilization by  
 XX preventing generation of fertile sperm.  
 XX Sequence 138 AA;

Query Match 65.4%; Score 696; DB 20; Length 138;  
 Best Local Similarity 69.7%; Pred. No. 1.8e-58;  
 Matches 138; Conservative 0; Mismatches 0; Indels 60; Gaps 1;  
 QY 1 MAKQPSVSSCEDREGROLQPAERPPQLRPGAPTSLOTPEOGNPEGNHGGEDSCPHGSP 60  
 DB 1 makqpsdvsscedregrlqpaerppqlrpgaptslqtpeq----- 41  
 QY 61 QGPLAPPASPCPFATRSPLFIEMRRSSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTSPP 120  
 DB 42 -----drspapmscdkstqtspp 60  
 QY 121 CQAFNHVLSAMSRQAEPADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQAAEDHPR 180  
 DB 61 cqafnhvlsamsmrqaeapadmripeiwiqlrrigdefnayyarrvflnnyqaedhpr 120  
 QY 181 MVILRLRYIVRLVWRMH 198  
 DB 121 mvilrllryivrlvwrmh 138  
 RESULT 5  
 AAW98165  
 ID AAW98165 standard; Protein; 138 AA.  
 XX  
 AC AAW98165;  
 XX  
 DT 05-JUL-1999 (first entry)  
 XX  
 DE Human Bim-L mutant S53p.  
 XX  
 KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; human; cancer; autoimmune disease; therapy;  
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
 KW cytotoxicity; agonist.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Binding-site 42..71 /note= "dynein light chain binding region, present  
 FT only in Bim-L, Bim-EL"  
 FT Misc-difference 53 /note= "Pro replaces Ser of wild-type Bim-L"  
 FT Region 94..102 /note= "BH3"  
 FT Region 123..137 /note= "hydrophobic region"  
 XX  
 XX WO9914321-A1.  
 XX 25-MAR-1999.  
 XX 17-SEP-1998; 98WO-AU00772.  
 XX 24-SEP-1997; 97AU-0009373.  
 XX 17-SEP-1997; 97AU-0009263.  
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 XX Puthalakath H, Strasser A;  
 XX WPI: 1999-244030/20.  
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 XX treatment  
 XX Claim 25; Page -; 145pp; English.  
 XX The present sequence is a S53p mutant of the long form (L) of

CC human Bim, or Bcl-2 interacting mediator of cell death (see also  
CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is  
CC capable of inducing cell death (apoptosis) and acts as a  
CC 'death-ligand' for certain members of the pro-survival Bcl-2  
CC family. Binding of the dynein light chain regulates the pro-apoptotic  
CC activity of Bim. Bim-S (see AAW98154), a splice variant which does  
CC not bind to dynein light chain, is a much more potent killer than  
CC either Bim-L or Bim-EL. The invention provides variants (see  
CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,  
CC couple or otherwise associate with a dynein light chain owing to  
CC amino acid addition, substitution and/or deletion within the region  
CC which binds the dynein light chain. The variants act as agonists  
CC of Bim function, interacting with anti-apoptotic molecules such as  
CC Bcl-2 to prevent their functional activity, thereby promoting  
CC apoptosis, and can be used e.g. in the treatment of cancer or to  
CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim  
CC is expressed in germ cells, modulating its expression or activity  
CC may be useful as a contraceptive or method of sterilization by  
CC preventing generation of fertile sperm.  
CC Note. The present sequence is not shown in the specification but  
CC is derived from the human Bim-L sequence given on page 100-101.  
XX  
XX Sequence 138 AA;  
SQ

Query Match 64.9%; Score 691; DB 20; Length 138;  
Best Local Similarity 69.2%; Pred. No. 5.4e-58;  
Matches 137; Conservative 0; Mismatches 1; Indels 60; Gaps 1;  
1;  
QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTEPGNHNHGGSDSCPHGSP 60  
DB 1 makqpsdvssecdregrlqpaerppqlrpgaptslqtpeq----- 41  
QY 61 QGPLAPPASGPPATRSPLFIEMRRSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120  
DB 42 -----drspapmscdkptqtsp 60  
QY 121 COAFNHYLSAMASMRQAEPADMRPEIWAQELRRIGDEFNAYYARRVFLNNYQAEDHPR 180  
DB 61 cqafnhylsamasmrqaeapadmrpeiwaqelrrigdefnayyarrvflnnyyqaadhr 120  
QY 181 MVILRLRLRYIVRLVWRMH 198  
DB 121 mvilrlrlryivrlvwrhm 138

RESULT 6  
AAW98166  
ID AAW98166 standard; Protein; 138 AA.  
XX  
XX AAW98166;  
XX  
XX 05-JUL-1999 (first entry)  
XX Human Bim-L mutant T54A.  
XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
KW cell cycle; human; cancer; autoimmune disease; therapy;  
KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
KW cytotoxicity; agonist.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Binding-site 42..71  
FT /note= "dynein light chain binding region, present  
FT Only in Bim-L, Bim-EL"  
FT Misc-difference 54  
FT /note= "Ala replaces Thr of wild-type Bim-L"  
FT Region 94..102  
FT /note= "BH3"  
FT Region 123..137

/note= "hydrophobic region"  
FT  
XX WO9914321-A1.  
PN  
XX  
XX 25-MAR-1999.  
PD  
XX  
XX 17-SEP-1998; 98WO-AU00772.  
PF  
XX  
XX 24-SEP-1997; 97AU-0009373.  
PR  
XX  
XX 17-SEP-1997; 97AU-0009263.  
PR  
XX  
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PA  
XX  
XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
PI Puthalakath H, Strasser A;  
XX  
XX WPI; 1999-244030/20.  
DR  
XX  
XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
PT treatment  
PT  
XX  
XX Claim 25; Page -; 145pp; English.  
PS  
XX  
XX The present sequence is a T54A mutant of the long form (L) of  
CC human Bim, or Bcl-2 interacting mediator of cell death (see also  
CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is  
CC capable of inducing cell death (apoptosis) and acts as a  
CC 'death-ligand' for certain members of the pro-survival Bcl-2  
CC family. Binding of the dynein light chain regulates the pro-apoptotic  
CC activity of Bim. Bim-S (see AAW98154), a splice variant which does  
CC not bind to dynein light chain, is a much more potent killer than  
CC either Bim-L or Bim-EL. The invention provides variants (see  
CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,  
CC couple or otherwise associate with a dynein light chain owing to  
CC amino acid addition, substitution and/or deletion within the region  
CC which binds the dynein light chain. The variants act as agonists  
CC of Bim function, interacting with anti-apoptotic molecules such as  
CC Bcl-2 to prevent their functional activity, thereby promoting  
CC apoptosis, and can be used e.g. in the treatment of cancer or to  
CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim  
CC is expressed in germ cells, modulating its expression or activity  
CC may be useful as a contraceptive or method of sterilization by  
CC preventing generation of fertile sperm.  
CC Note. The present sequence is not shown in the specification but  
CC is derived from the human Bim-L sequence given on page 100-101.  
XX  
XX Sequence 138 AA;  
SQ

Query Match 64.9%; Score 691; DB 20; Length 138;  
Best Local Similarity 69.2%; Pred. No. 5.4e-58;  
Matches 137; Conservative 0; Mismatches 1; Indels 60; Gaps 1;  
1;  
QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTEPGNHNHGGSDSCPHGSP 60  
DB 1 makqpsdvssecdregrlqpaerppqlrpgaptslqtpeq----- 41  
QY 61 QGPLAPPASGPPATRSPLFIEMRRSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120  
DB 42 -----drspapmscdktsqtsp 60  
QY 121 COAFNHYLSAMASMRQAEPADMRPEIWAQELRRIGDEFNAYYARRVFLNNYQAEDHPR 180  
DB 61 cqafnhylsamasmrqaeapadmrpeiwaqelrrigdefnayyarrvflnnyyqaadhr 120  
QY 181 MVILRLRLRYIVRLVWRMH 198  
DB 121 mvilrlrlryivrlvwrhm 138

RESULT 7  
AAW98168  
ID AAW98168 standard; Protein; 138 AA.

XX AC AAW98168;  
 XX XX  
 XX DT 05-JUL-1999 (first entry)  
 XX DE Human Bim-L mutant N65S.  
 XX XX  
 XX KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; human; cancer; autoimmune disease; therapy;  
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
 KW cytotoxicity; agonist.  
 XX XX  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX XX  
 XX FH Key Location/Qualifiers  
 XX Binding-site 42..71  
 XX FT /note= "dynein light chain binding region, present  
 FT only in Bim-L, Bim-EL"  
 FT  
 FT Misc-difference 65  
 FT /note= "Ser replaces Asn of wild-type Bim-L"  
 FT 94..102  
 FT /note= "BH3"  
 FT 123..137  
 FT /note= "hydrophobic region"  
 XX XX  
 XX PN WO9914321-A1.  
 XX XX  
 XX PD 25-MAR-1999.  
 XX XX  
 XX PF 17-SEP-1998; 98WO-AU000772.  
 XX XX  
 XX PR 24-SEP-1997; 97AU-0009373.  
 XX PR 17-SEP-1997; 97AU-0009263.  
 XX XX  
 XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX XX  
 XX PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX DR WPI; 1999-244030/20.  
 XX XX  
 XX XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 XX treatment  
 XX PS Claim 25; Page -; 145pp; English.  
 XX XX  
 XX CC The present sequence is a N65S mutant of the long form (L) of  
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also  
 CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is  
 CC capable of inducing cell death (apoptosis) and acts as a  
 CC 'death-ligand' for certain members of the pro-survival Bcl-2  
 CC family. Binding the dynein light chain regulates the pro-apoptotic  
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does  
 CC not bind to dynein light chain, is a much more potent killer than  
 CC either Bim-L or Bim-EL. The invention provides variants (see  
 CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,  
 CC couple or otherwise associate with a dynein light chain owing to  
 CC amino acid addition, substitution and/or deletion within the region  
 CC which binds the dynein light chain. The variants act as agonists  
 CC of Bim function, interacting with anti-apoptotic molecules such as  
 CC Bcl-2 to prevent their functional activity, thereby promoting  
 CC apoptosis, and can be used e.g. in the treatment of cancer or to  
 CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim  
 CC is expressed in germ cells, modulating its expression or activity  
 CC may be useful as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the human Bim-L sequence given on page 100-101.  
 XX XX  
 XX SQ Sequence 138 AA;

Query Match 64.9%; Score 691; DB 20; Length 138;  
 Best Local Similarity 69.2%; Pred. No. 5.4e-58;  
 Matches 137; Conservative 1; Mismatches 0; Indels 60; Gaps 1;  
 QY 1 MAKQPSDVSSCEDREGLOQPAERPPQLRPGAPTSLQTEPOGNEGNGEGDSCPHGSP 60  
 DB 1 makqpsdvsscedreglqlqaerppqlrpgaptslqtepq----- 41  
 QY 61 QGPLAPPASPGPFATRSPFLIFMRSSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120  
 DB 42 -----drspapmscdkstqtpsp 60  
 QY 121 COAFNHYLSAMASMRQAEPPADMRPEIWTAEQLRRTGDEFNAYYARRVFLNNYQAAEDHPR 180  
 DB 61 cqafshylsamasmrqaepadmreipeiwaqelrrigdefnayyarvflnnvqaaedhpr 120  
 QY 181 MVILRLRYIVRLVWRMH 198  
 DB 121 mvilrlryivrlvwrmh 138  
 RESULT 8  
 AAW98167  
 ID AAW98167 standard; Protein; 138 AA.  
 XX AC AAW98167;  
 XX XX  
 XX DT 05-JUL-1999 (first entry)  
 XX DE Human Bim-L mutant T54I.  
 XX KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; human; cancer; autoimmune disease; therapy;  
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
 KW cytotoxicity; agonist.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX XX  
 XX FH Key Location/Qualifiers  
 XX Binding-site 42..71  
 XX FT /note= "dynein light chain binding region, present  
 FT only in Bim-L, Bim-EL"  
 FT Misc-difference 54  
 FT /note= "Ile replaces Thr of wild-type Bim-L"  
 FT 94..102  
 FT /note= "BH3"  
 FT 123..137  
 FT /note= "hydrophobic region"  
 XX XX  
 XX PN WO9914321-A1.  
 XX PD 25-MAR-1999.  
 XX XX  
 XX PF 17-SEP-1998; 98WO-AU000772.  
 XX PR 24-SEP-1997; 97AU-0009373.  
 XX PR 17-SEP-1997; 97AU-0009263.  
 XX XX  
 XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 XX XX  
 XX PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX DR WPI; 1999-244030/20.  
 XX XX  
 XX XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 XX treatment  
 XX PS Claim 25; Page -; 145pp; English.  
 XX XX  
 XX CC The present sequence is a T54I mutant of the long form (L) of  
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also

CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is  
 CC capable of inducing cell death (apoptosis) and acts as a  
 CC 'death-ligand' for certain members of the pro-survival Bcl-2  
 CC family. Binding the dynein light chain regulates the pro-apoptotic  
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does  
 CC not bind to dynein light chain, is a much more potent killer than  
 CC either Bim-L or Bim-EL. The invention provides variants (see  
 CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,  
 CC couple or otherwise associate with a dynein light chain owing to  
 CC amino acid addition, substitution and/or deletion within the region  
 CC which binds the dynein light chain. The variants act as agonists  
 CC of Bim function, interacting with anti-apoptotic molecules such as  
 CC Bcl-2 to prevent their functional activity, thereby promoting  
 CC apoptosis, and can be used e.g. in the treatment of cancer or to  
 CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim  
 CC is expressed in germ cells, modulating its expression or activity  
 CC may be useful as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 CC Note. The present sequence is not shown in the specification but  
 CC is derived from the human Bim-L sequence given on page 100-101.  
 XX  
 SQ Sequence 138 AA;

Query Match 64.8%; Score 690; DB 20; Length 138;  
 Best Local Similarity 69.2%; Pred. No. 6.7e-58;  
 Matches 137; Conservative 0; Mismatches 1; Indels 60; Gaps 1;

QY 1 MAKQSDVSSDCRGRQLQPAERPPQLRPGAPTSLSQTEPQGNPEGNGEGDSCPHGSP 60  
 DB 1 makqsdvsdsccdrgrqlqpaerppqlrpgaptslqtpeq----- 41  
 QY 61 QGFLAPPASPGPFATRSPLFIEMRRSLLSRSSGYSFDTDRSPAPMSCDKSTQTPSP 120  
 DB 42 -----drspapmscdksiqtpsp 60  
 QY 121 COAFNHLSAMASMRQAEADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDHPR 180  
 DB 61 cqafnhlsamsmrqaepadmrpeiwiqaelrrigdefnayyarrvflnnyyqaedhpr 120  
 QY 181 MWILRLRYIVRLVWRMH 198  
 DB 121 mwilrlryivrlvwrmh 138

RESULT 9  
 ID AAW98155 standard; Protein; 140 AA.  
 AC AAW98155;  
 DT 05-JUL-1999 (first entry)  
 DE Murine Bcl-2 interacting mediator of cell death Bim-L isoform.  
 XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
 KW cell cycle; mouse; cancer; autoimmune disease;  
 KW degenerative disease; therapy; contraceptive; splice variant;  
 KW isoform.  
 XX Mus musculus.  
 XX Key Location/Qualifiers  
 FT Binding-site 42..71  
 FT /note= "dynein light chain binding region, present  
 FT only in Bim-L, Bim-EL"  
 FT Region 94..102  
 FT /note= "BH3"  
 FT Region 123..137  
 FT /note= "hydrophobic region"  
 XX WO9914321-A1.  
 PN  
 XX

PD 25-MAR-1999.  
 XX  
 PF 17-SEP-1998; 98WO-AU00772.  
 XX  
 PR 24-SEP-1997; 97AU-0009373.  
 PR 17-SEP-1997; 97AU-0009263.  
 XX  
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
 PI Puthalakath H, Strasser A;  
 XX  
 DR WPI; 1999-244030/20.  
 DR N-PSDB; AAX24994.  
 XX  
 DR New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
 PT treatment  
 XX  
 PS Claim 11; Page 95-96; 145pp; English.  
 XX  
 CC The present sequence is the long form (L) of murine Bim, or Bcl-2  
 CC interacting mediator of cell death, a novel member of the Bcl-2  
 CC family that is capable of inducing cell death (apoptosis) and which  
 CC acts as a 'death-ligand' for certain members of the pro-survival  
 CC Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology  
 CC region which it encompasses is BH3. It is the only BH3-only protein  
 CC for which splice variants exist. These result in the expression of  
 CC a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL (see AAW98154-56).  
 CC cDNAs encoding these murine Bim isoforms (see AAX24993-95) were  
 CC obtained from a T lymphoma cDNA library using human recombinant  
 CC Bcl-2 protein. Human Bim-L and Bim-EL isoforms have also been  
 CC identified (see AAW98157-58). Binding the dynein light chain was shown  
 CC to regulate the pro-apoptotic activity of Bim. Bim-S, the splice  
 CC variant which does not bind to dynein light chain, is a much more  
 CC potent killer than either Bim-L or Bim-EL. The invention provides  
 CC variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that  
 CC cannot bind, couple or otherwise associate with a dynein light  
 CC chain. The identification of Bim permits the identification and  
 CC rational design of a range of products for use in therapy.  
 CC diagnosis, antibody generation and involving modulation of  
 CC physiological cell death. These therapeutic molecules may act  
 CC as either antagonists or agonists of Bim's function and will be  
 CC useful in cancer, autoimmune or degenerative disease therapy.  
 CC Increased Bim expression or Bim activity is useful, e.g. for  
 CC treatment or prophylaxis in conditions such as cancer and deletion  
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim  
 CC expression of Bim activity is useful in regulating inhibition or  
 CC prevention of cell death or degeneration such as under cytotoxic  
 CC conditions during e.g. gamma-irradiation and chemotherapy or during  
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,  
 CC hypoxia, degenerative diseases or for prolonging the survival of  
 CC cells being transplanted for treatment of disease. Since Bim is  
 CC expressed in germ cells, modulating Bim expression or Bim activity  
 CC is useful, e.g. as a contraceptive or method of sterilization by  
 CC preventing generation of fertile sperm.  
 XX  
 SQ Sequence 140 AA;

Query Match 55.3%; Score 589; DB 20; Length 140;  
 Best Local Similarity 60.0%; Pred. No. 2.7e-48;  
 Matches 120; Conservative 8; Mismatches 10; Indels 62; Gaps 2;

QY 1 MAKQSDVSSDCRGRQLQPAERPPQLRPGAPTSLSQTEPQGNPEGNGEGDSCPHGSP 60  
 DB 1 makqsdvsdsccdrgrqlqpaerppqlrpgaptslqtpeq----- 41  
 QY 61 QGFLAPPASPGPFATRSPLFIEMRRSLLSRSSGYSFDTDRSPAPMSCDKSTQTPSP 120  
 DB 42 -----drspapmscdksiqtpsp 60  
 QY 121 COAFNHLSAMASMRQAEADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 178  
 XX

Db 61 cqafnhylsamarsirgseepdlrpeiriaeqlrrrigdefnetytrrvfandyreaedh 120

Qy 179 PRMVLRLRLRYIVRLVWRMH 198  
l:||||:||||:| ||||| |

Db 121 pqmvilqlrlfrlrvwrh 140

RESULT 10  
AAW98160  
ID AAW98160 standard; Protein: 140 AA.  
XX AC AAW98160;  
XX 05-JUL-1999 (first entry)  
XX Murine Bim-L mutant S53P.  
XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
KW cell cycle; mouse; cancer; autoimmune disease; therapy;  
KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
KW cytotoxicity; agonist.  
XX Mus musculus.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Binding-site 42..71 /note= "dynein light chain binding region, present  
FT only in Bim-L, Bim-EL"  
FT Misc-difference 53 /note= "Pro replaces Ser of wild-type Bim-L"  
FT Region 94..102 /note= "BH3"  
FT Region 123..137 /note= "hydrophobic region"  
XX WO9914321-A1.  
XX 25-MAR-1999.  
XX 17-SEP-1998; 98WO-AU00772.  
XX 24-SEP-1997; 97AU-0009373.  
XX 17-SEP-1997; 97AU-0009263.  
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
PI Puthalakath H, Strasser A;  
XX WPI; 1999-244030/20.  
XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer  
PT treatment  
XX Claim 25; Page -: 145pp; English.  
XX The present sequence is a S53p mutant of the long form (L) of  
CC murine Bim, or Bcl-2 interacting mediator of cell death (see also  
CC AAW98155), a novel member of the Bcl-2 family. Native Bim-L is  
CC capable of inducing cell death (apoptosis) and acts as a  
CC 'death-ligand' for certain members of the pro-survival Bcl-2  
CC family. Binding the dynein light chain regulates the pro-apoptotic  
CC activity of Bim. Bim-S (see AAW98154), a splice variant which does  
CC not bind to dynein light chain, is a much more potent killer than  
CC either Bim-L or Bim-EL. The invention provides variants (see  
CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,  
CC couple or otherwise associate with a dynein light chain owing to  
CC amino acid addition, substitution and/or deletion within the region  
CC which binds the dynein light chain. The variants act as agonists  
CC of Bim function, interacting with anti-apoptotic molecules such as  
CC Bcl-2 to prevent their functional activity, thereby promoting  
CC apoptosis, and can be used e.g. in the treatment of cancer or to

CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim  
CC is expressed in germ cells, modulating its expression or activity  
CC may be useful as a contraceptive or method of sterilization by  
CC preventing generation of fertile sperm.  
CC Note. The present sequence is not shown in the specification but  
CC is derived from the murine Bim-L sequence given on page 95-96.  
XX XX Sequence 140 AA;  
SQ

Query Match 54.8%; Score 584; DB 20; Length 140;  
Best Local Similarity 59.5%; Pred. No. 8.1e-48;  
Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;  
Qy 1 MAKQPSDVSSCEDREGRLQPAERPPQLRPGCAPTSLOTPEQCNGNHRGGSDSCPHGSP 60  
|||||:|||||:| ||||| |  
Db 1 makpsdvssecdreggqlqpaerppqlrpgaptslqteq----- 41  
Qy 61 QGPIAPPASGPPFATRSPLFIEMRRSSLLSSSGSYFSFDTRSPAPMSCDKSTQTPSP 120  
|||||:|||||:| ||||| |  
Db 42 -----drspapmscdkptqtpsp 60  
Qy 121 CQAFNHYLSAMASMRQA--EPADMRPEITWIAQELRLRIGDEFNAYYARRVFLNNYQAAEDH 178  
|||||:|||||:| ||||| |  
Db 61 cqafnhylsamarsirgseepdlrpeiriaeqlrrrigdefnetytrrvfandyreaedh 120  
Qy 179 PRMVLRLRLRYIVRLVWRMH 198  
l:||||:||||:| ||||| |  
Db 121 pqmvilqlrlfrlrvwrh 140

RESULT 11  
AAW98161  
ID AAW98161 standard; Protein: 140 AA.  
XX AC AAW98161;  
XX 05-JUL-1999 (first entry)  
XX Murine Bim-L mutant T54A.  
XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;  
KW cell cycle; mouse; cancer; autoimmune disease; therapy;  
KW contraceptive; splice variant; isoform; mutant; dynein light chain;  
KW cytotoxicity; agonist.  
XX Mus musculus.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Binding-site 42..71 /note= "dynein light chain binding region, present  
FT only in Bim-L, Bim-EL"  
FT Misc-difference 54 /note= "Ala replaces Thr of wild-type Bim-L"  
FT Region 94..102 /note= "BH3"  
FT Region 123..137 /note= "hydrophobic region"  
XX WO9914321-A1.  
XX 25-MAR-1999.  
XX 17-SEP-1998; 98WO-AU00772.  
XX 24-SEP-1997; 97AU-0009373.  
XX 17-SEP-1997; 97AU-0009263.  
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;  
PI Puthalakath H, Strasser A;



```

XX DR WPI; 1999-244030/20.
XX PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
XX PT treatment
XX PS Claim 25; Page -: 145pp; English.
XX CC The present sequence is a T54A mutant of the long form (L) of
XX CC murine Bim, or Bcl-2 interacting mediator of cell death (see also
XX CC AAW98155), a novel member of the Bcl-2 family. Native Bim-L is
XX CC capable of inducing cell death (apoptosis) and acts as a
XX CC 'death-ligand' for certain members of the pro-survival Bcl-2
XX CC family. Binding the dynein light chain regulates the pro-apoptotic
XX CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
XX CC not bind to dynein light chain, is a much more potent killer than
XX CC either Bim-L or Bim-EL. The invention provides variants (see
XX CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
XX CC couple or otherwise associate with a dynein light chain owing to
XX CC amino acid addition, substitution and/or deletion within the region
XX CC which binds the dynein light chain. The variants act as agonists
XX CC of Bim function, interacting with anti-apoptotic molecules such as
XX CC Bcl-2 to prevent their functional activity, thereby promoting
XX CC apoptosis, and can be used e.g. in the treatment of cancer or to
XX CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
XX CC is expressed in germ cells, modulating its expression or activity
XX CC may be useful as a contraceptive or method of sterilization by
XX CC preventing generation of fertile sperm.
XX CC Note. The present sequence is not shown in the specification but
XX CC is derived from the murine Bim-L sequence given on page 95-96.
XX SQ Sequence 140 AA;

Query Match 54.8%; Score 584; DB 20; Length 140;
Best Local Similarity 59.5%; Pred. No. 8.1e-48;
Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;

QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTPEQGNPEGNHGEGDCSPHSP 60
DB 1 makqpsdvsecdreggqlqpaerppqlrpgaptslqtepq----- 41

QY 61 QGPLAPPASGPPATRSPLIFMRRSLLSRSSGYFSFDTDRSPAPMSCKSTQTPSP 120
DB 42 -----drspapmscdksaqtsp 60

QY 121 COAFNHVLSAMASMRQA--EPADMRPEIWTAEQLRIGDEFNAYYARRVFLNNYQAEADH 178
DB 61 cqafnhylsamasirsdqeepdlrpeiriaqelrrigdefnetyrrrvfandyreadh 120

QY 179 PRMVLRLRYIVRLVWRMH 198
DB 121 pgmvlrlrlryivrlvwrwh 140

RESULT 12
AAW98163
ID AAW98163 standard; Protein; 140 AA.
AC AAW98163;
XX 05-JUL-1999 (first entry)
XX Murine Bim-L mutant N65S.
XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
XX cell cycle; mouse; cancer; autoimmune disease; therapy.
XX KW contraceptive; splice variant; isoform; mutant; dynein light chain;
XX KW cytotoxicity; agonist.
XX OS Mus musculus.
XX OS Synthetic.
XX

```

```

FH Key Location/Qualifiers
FT Binding-site 42..71
FT /note= "dynein light chain binding region, present
FT only in Bim-L, Bim-EL"
FT Misc-difference 65
FT /note= "Ser replaces Asn of wild-type Bim-L"
FT Region 94..102
FT /note= "BH3"
FT Region 123..137
FT /note= "hydrophobic region"
XX WO9914321-A1.
XX 25-MAR-1999.
XX 17-SEP-1998; 98WO-AU00772.
XX 24-SEP-1997; 97AU-0009373.
XX 17-SEP-1997; 97AU-0009263.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX Adams J. Cory S. Huang DCS, O'Connor L, O'Reilly L;
XX Puthalakath H, Strasser A;
XX WPI; 1999-244030/20.
XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
XX treatment
XX Claim 25; Page -: 145pp; English.
XX The present sequence is a N65S mutant of the long form (L) of
XX murine Bim, or Bcl-2 interacting mediator of cell death (see also
XX AAW98155), a novel member of the Bcl-2 family. Native Bim-L is
XX capable of inducing cell death (apoptosis) and acts as a
XX 'death-ligand' for certain members of the pro-survival Bcl-2
XX family. Binding the dynein light chain regulates the pro-apoptotic
XX activity of Bim. Bim-S (see AAW98154), a splice variant which does
XX not bind to dynein light chain, is a much more potent killer than
XX either Bim-L or Bim-EL. The invention provides variants (see
XX AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
XX couple or otherwise associate with a dynein light chain owing to
XX amino acid addition, substitution and/or deletion within the region
XX which binds the dynein light chain. The variants act as agonists
XX of Bim function, interacting with anti-apoptotic molecules such as
XX Bcl-2 to prevent their functional activity, thereby promoting
XX apoptosis, and can be used e.g. in the treatment of cancer or to
XX delete autoreactive lymphocytes in autoimmune diseases. Since Bim
XX is expressed in germ cells, modulating its expression or activity
XX may be useful as a contraceptive or method of sterilization by
XX preventing generation of fertile sperm.
XX Note. The present sequence is not shown in the specification but
XX is derived from the murine Bim-L sequence given on page 95-96.
XX SQ Sequence 140 AA;

Query Match 54.8%; Score 584; DB 20; Length 140;
Best Local Similarity 59.5%; Pred. No. 8.1e-48;
Matches 119; Conservative 9; Mismatches 10; Indels 62; Gaps 2;

QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTPEQGNPEGNHGEGDCSPHSP 60
DB 1 makqpsdvsecdreggqlqpaerppqlrpgaptslqtepq----- 41

QY 61 QGPLAPPASGPPATRSPLIFMRRSLLSRSSGYFSFDTDRSPAPMSCKSTQTPSP 120
DB 42 -----drspapmscdksaqtsp 60

QY 121 COAFNHVLSAMASMRQA--EPADMRPEIWTAEQLRIGDEFNAYYARRVFLNNYQAEADH 178
DB 61 cqafnhylsamasirsdqeepdlrpeiriaqelrrigdefnetyrrrvfandyreadh 120

```



WPI; 1999-244030/20.

New isolated member of the Bcl-2 family, Bim used in, e.g. cancer treatment

Claim 25; Page -: 145pp; English.

The present sequence is a D51G mutant of the long form (L) of murine Bim, or Bcl-2 interacting mediator of cell death (see also AAW98155), a novel member of the Bcl-2 family. Native Bim-L is capable of inducing cell death (apoptosis) and acts as a 'death-ligand' for certain members of the pro-survival Bcl-2 family. Binding the dynein light chain regulates the pro-apoptotic activity of Bim. Bim-S (see AAW98154), a splice variant which does not bind to dynein light chain, is a much more potent killer than either Bim-L or Bim-EL. The invention provides variants (see AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind, couple or otherwise associate with a dynein light chain owing to amino acid addition, substitution and/or deletion within the region which binds the dynein light chain. The variants act as agonists of Bim function, interacting with anti-apoptotic molecules such as Bcl-2 to prevent their functional activity, thereby promoting apoptosis, and can be used e.g. in the treatment of cancer or to delete autoreactive lymphocytes in autoimmune diseases. Since Bim is expressed in germ cells, modulating its expression or activity may be useful as a contraceptive or method of sterilization by preventing generation of fertile sperm.

Note. The present sequence is not shown in the specification but is derived from the murine Bim-L sequence given on page 95-96.

Sequence 140 AA;

Query Match 54.6%; Score 582; DB 20; Length 140;  
 Best Local Similarity 59.5%; Pred. No. 1.3e-47;  
 Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSECDEGRQLQPAERPPOLRPGAPTSLOTPEQGNPEGNHGSGDSCPHGSP 60  
 |||||  
 Db 1 makqpsdvssecdegrgqlpaerppolrpgaptslqtpeq----- 41

QY 61 QGPLAPPASGPPATRSPLPFIMRRSLLSRSSGYSFDFDRSPAPMSCDKSTQTSPSP 120  
 |||||  
 Db 42 -----drspapmscgkstqtpssp 60

QY 121 CQAFNHVLSAMASMRQA--EPADMRIPIWIAQELRRIGDEFNAYYARRVFLNNYQAAEDH 178  
 |||||  
 Db 61 cqafnhvlsamasirsdqeedrpeiraiqelrrigdefnetytrrvfandyreaedh 120  
 |||||  
 QY 179 PRMVIILRLRYIVRLVWRMH 198  
 |||||  
 Db 121 pqmviqlrlrfiflvrwrh 140

RESULT 15  
 ABG01163  
 ID ABG01163 standard; Protein; 92 AA.  
 XX ABG01163;  
 XX  
 XX 13-FEB-2002 (first entry)  
 XX Novel human diagnostic protein #1154.  
 XX  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200175067-A2.  
 XX  
 XX 11-OCT-2001.  
 XX



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 09:57:22 ; Search time 13.18 Seconds  
(without alignments)  
366.940 Million cell updates/sec

Title: US-09-508-832-10  
Perfect score: 1065  
Sequence: 1 MAKQPSDVSSECDREGRLQ.....PRMVLRLRYIVRLVRMH 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111.5	10.5	886	US-08-474-379C-65	Sequence 65, Appl
2	111.5	10.5	886	US-09-146-249A-65	Sequence 65, Appl
3	111.5	10.5	886	US-08-206-188B-65	Sequence 65, Appl
4	108.5	10.2	885	US-08-577-492-33	Sequence 33, Appl
5	108.5	10.2	885	US-09-079-630-33	Sequence 33, Appl
6	99	9.3	228	US-09-219-849-38	Sequence 38, Appl
7	99	9.3	1057	US-08-931-820-4	Sequence 4, Appl
8	99	9.3	1078	US-08-963-825-21	Sequence 21, Appl
9	99	9.3	1078	US-09-570-573-21	Sequence 21, Appl
10	99	9.3	1078	US-09-548-608-21	Sequence 21, Appl
11	94.5	8.9	280	US-09-247-155-178	Sequence 178, App
12	94.5	8.9	1185	US-09-041-886-23	Sequence 23, Appl
13	94	8.8	580	US-08-906-865-1	Sequence 1, Appl
14	94	8.8	627	US-08-466-589-6	Sequence 6, Appl
15	94	8.8	627	US-08-700-636-6	Sequence 6, Appl
16	94	8.8	627	US-08-467-574-6	Sequence 6, Appl
17	94	8.8	627	US-09-217-345-6	Sequence 6, Appl
18	93.5	8.8	434	US-08-710-249-4	Sequence 4, Appl
19	93.5	8.8	434	US-09-220-157A-4	Sequence 4, Appl
20	93.5	8.8	902	US-08-396-479B-6	Sequence 6, Appl
21	93.5	8.8	902	US-08-818-823-6	Sequence 6, Appl
22	92.5	8.7	245	US-09-140-804-4	Sequence 4, Appl
23	92.5	8.7	829	US-08-642-255-132	Sequence 132, App
24	92.5	8.7	829	US-08-397-633A-53	Sequence 53, Appl
25	92	8.6	761	US-08-707-237A-84	Sequence 84, Appl
26	92	8.6	762	US-08-642-255-114	Sequence 114, App
27	92	8.6	762	US-08-642-255-120	Sequence 120, App

28	92	8.6	762	1	US-08-397-633A-26	Sequence 26, Appl
29	92	8.6	762	1	US-08-397-633A-31	Sequence 31, Appl
30	91.5	8.6	1319	2	US-08-290-731C-2	Sequence 2, Appl
31	91.5	8.6	1336	2	US-08-290-731C-6	Sequence 6, Appl
32	90	8.5	1323	1	US-08-026-138E-4	Sequence 4, Appl
33	90	8.5	1706	2	US-08-459-568-2	Sequence 2, Appl
34	90	8.5	1706	2	US-08-399-411-2	Sequence 2, Appl
35	90	8.5	1706	3	US-08-516-859A-2	Sequence 2, Appl
36	89.5	8.4	123	4	US-09-247-155-91	Sequence 91, Appl
37	89	8.4	1719	2	US-08-459-568-4	Sequence 4, Appl
38	89	8.4	1719	2	US-08-399-411-4	Sequence 4, Appl
39	89	8.4	1719	3	US-08-516-859A-4	Sequence 4, Appl
40	88.5	8.3	1333	3	US-09-356-952-2	Sequence 2, Appl
41	88	8.3	357	1	US-07-609-716-66	Sequence 66, Appl
42	88	8.3	357	1	US-08-642-255-33	Sequence 33, Appl
43	88	8.3	357	4	US-08-475-411A-66	Sequence 66, Appl
44	88	8.3	357	4	US-08-478-029A-66	Sequence 66, Appl
45	87.5	8.2	2337	3	US-08-713-118-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-474-379C-65  
; Sequence 65, Application US/08474379C  
; Patent No. 5977305  
; GENERAL INFORMATION:  
; APPLICANT: Wigler, Michael H.  
; APPLICANT: Colicelli, John J.  
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED  
; TITLE OF INVENTION: PROCESSES  
; NUMBER OF SEQUENCES: 88  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/474,379C  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/511,715  
; FILING DATE: 20-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/206,188  
; FILING DATE: 01-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/688,352  
; FILING DATE: 19-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Crough, David W.  
; REGISTRATION NUMBER: 36,107  
; REFERENCE/DOCKET NUMBER: 27866/32771  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 886 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-474-379C-65

Query Match 10.5%; Score 111.5; DB 2; Length 886;  
Best Local Similarity 27.5%; Pred. No. 0.0046;  
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;  
QY 6 SDVSECDREGROLQAPERPQL----RPGAPTSLOTEPQ-----GNPEGNHGG----- 50  
DB 49 SDSAERAERQPHRIERADAMTSDRPGLTRTMSWPSFHGTGTGSGGAGGSSRRF 108  
QY 51 EGDSCPHGSP-QGPLAPPSPGPATRSPLFIEMRRSLLSRSSSGYFSDTDRSPAPMS 109  
DB 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162  
QY 110 CDKSTQTPSPCCQAFNHYLSAMASMRQAEADMPMEIWIQAELRRIGDEFN 160  
DB 163 RN-----SSVTSEAHAEADLIVTPFAQVLASLRSVRNFS 196  
RESULT 2  
US-09-146-249A-65  
; Sequence 65, Application US/09146249A  
; Patent No. 6069240  
; GENERAL INFORMATION:  
; APPLICANT: Wigler, Michael H.  
; APPLICANT: Colicelli, John J.  
; TITLE OF INVENTION: Cloning by Complementation and Related  
; TITLE OF INVENTION: Processes  
; NUMBER OF SEQUENCES: 85  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/146,249A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/511,715  
; FILING DATE: 20-APR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clough, David W.  
; REGISTRATION NUMBER: 36,107  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 886 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-146-249A-65  
Query Match 10.5%; Score 111.5; DB 3; Length 886;  
Best Local Similarity 27.5%; Pred. No. 0.0046;  
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;  
QY 6 SDVSECDREGROLQAPERPQL----RPGAPTSLOTEPQ-----GNPEGNHGG----- 50  
DB 49 SDSAERAERQPHRIERADAMTSDRPGLTRTMSWPSFHGTGTGSGGAGGSSRRF 108  
QY 51 EGDSCPHGSP-QGPLAPPSPGPATRSPLFIEMRRSLLSRSSSGYFSDTDRSPAPMS 109  
DB 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162  
QY 110 CDKSTQTPSPCCQAFNHYLSAMASMRQAEADMPMEIWIQAELRRIGDEFN 160  
DB 163 RN-----SSVTSEAHAEADLIVTPFAQVLASLRSVRNFS 196

Db 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162  
QY 110 CDKSTQTPSPCCQAFNHYLSAMASMRQAEADMPMEIWIQAELRRIGDEFN 160  
Db 163 RN-----SSVTSEAHAEADLIVTPFAQVLASLRSVRNFS 196  
RESULT 3  
US-08-206-188B-65  
; Sequence 65, Application US/08206188B  
; Patent No. 6100025  
; GENERAL INFORMATION:  
; APPLICANT: Wigler, Michael H.  
; APPLICANT: Colicelli, John J.  
; TITLE OF INVENTION: Cloning by Complementation and Related  
; TITLE OF INVENTION: Processes  
; NUMBER OF SEQUENCES: 84  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/206,188B  
; FILING DATE: 01-MAR-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/511,715  
; FILING DATE: 20-APR-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clough, David W.  
; REGISTRATION NUMBER: 36107  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 886 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-206-188B-65  
Query Match 10.5%; Score 111.5; DB 3; Length 886;  
Best Local Similarity 27.5%; Pred. No. 0.0046;  
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;  
QY 6 SDVSECDREGROLQAPERPQL----RPGAPTSLOTEPQ-----GNPEGNHGG----- 50  
DB 49 SDSAERAERQPHRIERADAMTSDRPGLTRTMSWPSFHGTGTGSGGAGGSSRRF 108  
QY 51 EGDSCPHGSP-QGPLAPPSPGPATRSPLFIEMRRSLLSRSSSGYFSDTDRSPAPMS 109  
Db 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162  
QY 110 CDKSTQTPSPCCQAFNHYLSAMASMRQAEADMPMEIWIQAELRRIGDEFN 160  
Db 163 RN-----SSVTSEAHAEADLIVTPFAQVLASLRSVRNFS 196  
RESULT 4  
US-08-577-492-33  
; Sequence 33, Application US/08577492  
; Patent No. 5851784

```

; GENERAL INFORMATION:
; APPLICANT: Owens, Raymond John
; APPLICANT: Perry, Martin John
; APPLICANT: Lumb, Simon Mark
; TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND
; TITLE OF INVENTION: ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5851784ris
; STREET: One liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/577,492
; FILING DATE: 22-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9426227.6
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9512996.1
; FILING DATE: 26-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cherry, David A.
; REGISTRATION NUMBER: 35,099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; US-08-577-492-33

Query Match 10.2%; Score 108.5; DB 2; Length 885;
Best Local Similarity 27.1%; Pred. No. 0.0092;
Matches 46; Conservative 18; Mismatches 69; Indels 37; Gaps

QY 6 SDVSECDREGQLQPAERPPQL-----RPCAPTSLOTPE---OGNPEGNHGGEGDS---- 54
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 49 SDSAEAEERQPHRPIERADAMDTSDRPLGRTIRMSWSPFSFHGTGSGGAGGSSRRF 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 55 -CPHGSPOG--PLAPPASPGP-FATRSPLFIEMRRSSLLSRSSSGYFSFDTDRSPAPWSC 11
   : : : : : || : || : || : || : || : || : || : || : || : || : ||
Db 109 EAENGTSAGRSPLDPMWTSFGPLVLHAGATSORRESFLYRSDDY-----DMSPTMSR 160
   : : : : : || : || : || : || : || : || : || : || : || : || : ||
Db 164 N-----SSVTSEAHAEEDLIVTPFAQVLASLRSVRSNFs 196

RESULT 5
US-09-079-630-33
; Sequence 33, Application US/09079630
; Patent No. 6291199
; GENERAL INFORMATION:
; APPLICANT: Owens, Raymond John
; APPLICANT: Perry, Martin John
; APPLICANT: Lumb, Simon Mark
; TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND
; TITLE OF INVENTION: ITS PRODUCTION AND USE
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6291199ris

```

```
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/079,630
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/577,492
; FILING DATE: 22-DEC-1995
; APPLICATION NUMBER: GB 9426227.6
; FILING DATE: 23-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9512996.1
; FILING DATE: 26-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Cherry, David A.
; REGISTRATION NUMBER: 35,099
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-09-079-630-33

Query Match          10.2%; Score 108.5; DB 4; Length 885;
Best Local Similarity 27.1%; Pred. No. 0.0092;
Matches 46; Conservative 18; Mismatches 69; Indels 37; Gaps

QY   6 SDVSSCEDRGLOPAERPQL-----RCAPITSLTEP---QGNPEGNHGEGDS---- 54
    || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|
Db   49 SDSAERARERQRPHRTIERADAMDTSDRPLRLTRMSPWPSFHGTCTGGAGGGSSRRF 108
    ..:::||:|||||::||:||||:||||:||||:||||:||||:||||:||||:|
QY   55 -CPHGSPQG--PLAPPASPG-PATRSPLFIEMRSSLSLSSSGVFSDTDRSPAPMSC 110
    :::|:|:|||||::||:||||:||||:||||:||||:||||:||||:||||:|
Db   109 EAENGTSAGSPLDPTSFPSGLVLHAGATSORRESFLYRSDSY-----DMSPKTMR 163
      |:|:|:|||||::||:||||:||||:||||:||||:||||:||||:||||:|
QY   111 DKSTQTTPPPCOAFNHYLSAMASMQAEPADMPEIWIQAELRRIGDEFN 160
       :||:|:|||||::||:||||:||||:||||:||||:||||:||||:||||:|
Db   164 N-----SSVTSEAHAEIDLIVTFPAQVLASLRVSRSNF'S 196
        :||:|:|||||::||:||||:||||:||||:||||:||||:||||:||||:|

RESULT               6
US-09-219-849-38
; Sequence 38, Application US/09219849
; Patent No. 6150081
; GENERAL INFORMATION:
; APPLICANT: VAN HEERDE, GEORGE V.
; APPLICANT: VAN RIJN, ALEXIS C.
; APPLICANT: BOUWSTRAJAN B.
; APPLICANT: DE WOLF, FREDERIK A.
; APPLICANT: MOOBROEK, ANDREAS
; APPLICANT: WERTEN, MARC W.T.
; APPLICANT: WIND, RICHELE D.
; APPLICANT: VAN DEN BOSCH, TANJA J.
; TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
; TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
; TITLE OF INVENTION: PREPARATION THEREOF
; FILE REFERENCE: 2728-2
; CURRENT APPLICATION NUMBER: US/09/219,849
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 50
```

[illegible]

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 38

; LENGTH: 228

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: amino acid sequence

US-09-219-849-38

Query Match

Best Local Similarity 35.4%; Score 99; DB 4; Length 228;

Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;

QY 5 PSDVSSCEDREGROLPAPER---PQLRPGAP-TSLQTEPGNP-----EGNHGEG 52

Db 66 PQIGKSGKPGAGSHNGERPPGQGLPGQGTAGPRDGNPGSDGQPGRDGSPGKG 125

QY 53 DSCPHGSPQGPPLAP--PASPGP 72

Db 126 DRGENSGPAGPAGHPGPPGP 147

RESULT 7

US-08-931-820-4

; Sequence 4, Application US/08931820

; Patent No. 6010863

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Assay for collagen degradation

; NUMBER OF SEQUENCES: 4

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/931,820

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 96202596.1

; FILING DATE:

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1057 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; TISSUE TYPE: Collagen type III

; FEATURE:

; NAME/KEY: Modified-site

; LOCATION: 1055

; OTHER INFORMATION: /label= Modified

; OTHER INFORMATION: /note= "Ala may be Pro"

US-08-931-820-4

Query Match

Best Local Similarity 27.8%; Score 99; DB 3; Length 1057;

Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCEDREGROLPAPERPPQLRPGAPTSLOT--EP--QGPN-----EGNHGEG 52

Db 825 PQGVKSGKPGANGLSGERPPGQGLPGLAGTAGPRDGNPGSDGLPCRDGSPGKG 884

QY 53 DSCPHGSPQGPPLAP--PASPGPFATRSPLFTFMRRSSILSRSSSGYFSDTDRSPAPMSC 110

Db 111 DKSTQTPSPPCQA 123

Db 885 DRGENSGPAGPAGHPGPPGPV-----GPAGKSG 914

QY 111 DKSTQTPSPPCQA 123

Db 915 DRGENSGPAGPAGA 927

RESULT 8

US-08-963-825-21

; Sequence 21, Application US/08963825

; Patent No. 6110689

; GENERAL INFORMATION:

; APPLICANT: Qvist, Per

; APPLICANT: Bonde, Martin

; TITLE OF INVENTION: A Method for Assaying Collagen Fragments

; TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the

; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of

; TITLE OF INVENTION: Disorders Associated with the Metabolism of

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby PC

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/963,825

; FILING DATE:

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/187,319

; FILING DATE: 21-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Gogoris, Adda C

; REGISTRATION NUMBER: 29,714

; REFERENCE/DOCKET NUMBER: 4305/08701

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-527-7700

; TELEFAX: 212-753-6237

; TELEX: 236687

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1078 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; IMMEDIATE SOURCE:

; CLONE: COLLAGEN ALPHA 1 (III)

US-08-963-825-21

Query Match

Best Local Similarity 27.8%; Score 99; DB 3; Length 1078;

Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCEDREGROLPAPERPPQLRPGAPTSLOT--EP--QGPN-----EGNHGEG 52

Db 826 PQGVKSGKPGANGLSGERPPGQGLPGLAGTAGPRDGNPGSDGLPCRDGSPGKG 885

QY 53 DSCPHGSPQGPPLAP--PASPGPFATRSPLFTFMRRSSILSRSSSGYFSDTDRSPAPMSC 110

Db 886 DRGENSGPAGPAGHPGPPGPV-----GPAGKSG 915

QY 111 DKSTQTPSPPCQA 123









Qy 119 P-PCQ 122  
I II:  
Db 446 PGPCR 450

Search completed: August 16, 2002, 10:01:20  
Job time: 238 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 09:59:27 ; Search time 19.86 Seconds  
(without alignments)  
957.990 Million cell updates/sec

Title: US-09-508-832-10

Perfect score: 1065

Sequence: 1 MAKQPSDVSECDREGRLQ.....PRMVILRLRYIVRLWRMH 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111.5	10.5	886	2 A54442	3',5'-cyclic-nucle
2	110.5	10.4	472	2 T03169	probable glycoprot
3	105.5	9.9	574	2 T43556	Wiskott-Aldrich sy
4	105.5	9.9	574	2 T38819	wiskott-aldrich sy
5	103.5	9.7	556	2 T42100	serine/threonine p
6	103.5	9.7	556	2 T36502	serine/threonine p
7	102.5	9.6	1332	1 I48314	homeotic protein C
8	100	9.4	99	1 CGEHS	collagen-like prot
9	100	9.4	627	2 J24021	nicotinic acetylch
10	99	9.3	636	2 S41067	collagen alpha 1(I
11	99	9.3	1464	2 S59856	collagen alpha 1(I
12	99	9.3	1466	1 CGHU7L	collagen alpha 1(I
13	97.5	9.2	289	2 T26812	hypothetical prote
14	97.5	9.2	443	2 I49140	p62 ras-GAP associ
15	97	9.1	852	2 T10811	channel associated
16	96	9.0	870	2 G01974	channel associated
17	96	9.0	1315	2 A56101	collagen alpha 1(X
18	96	9.0	1774	2 B56101	collagen alpha 1(X
19	95.5	9.0	172	2 D41132	collagen-related p
20	95.5	9.0	210	2 B44984	collagen - nematod
21	95	8.9	298	2 J21448	collagen col-34 -
22	95	8.9	844	2 I53865	phosphodiesterase
23	94.5	8.9	573	2 B70726	probable secD - My
24	94.5	8.9	628	2 JQ0110	hypothetical 69K p
25	94.5	8.9	684	2 A53019	collagen alpha 1(X
26	94.5	8.9	1184	2 G01763	atrophin-1 - human
27	94.5	8.9	1184	2 S50832	atrophin-1 - human
28	94.5	8.9	1234	2 A34911	band 3-related pro
29	94	8.8	268	2 H84684	En/Spm-like transp

30	94	8.8	278	2 B39066	proline-rich prote
31	94	8.8	1460	1 EDBEIF	immediate-early pr
32	93.5	8.8	228	2 A44982	collagen UCO11 - p
33	93	8.7	1400	2 T52359	hypothetical prote
34	92.5	8.7	245	1 C1HUQA	complement subcomp
35	92.5	8.7	1446	1 A45344	immediate-early pr
36	92	8.6	464	2 S59513	collagen II A1 pro
37	92	8.6	589	2 A5968	folate binding prote
38	92	8.6	1172	2 T00065	hypothetical prote
39	91.5	8.6	310	1 P1HUSD	salivary proline-r
40	91.5	8.6	473	2 S50755	hypothetical prote
41	91.5	8.6	730	2 A36226	collagen alpha 1 c
42	91.5	8.6	744	2 T35192	probable ABC trans
43	91.5	8.6	1133	2 T12529	hypothetical prote
44	91.5	8.6	1336	2 S25716	Ras guanine nucleo
45	91	8.5	260	2 S22373	proline-rich prote

## ALIGNMENTS

RESULT 1  
A54442  
3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) 4A, cAMP-specific, long splic  
N:Contains: 3',5'-cyclic AMP phosphodiesterase HPDE4A6 splice form  
C:Species: Homo sapiens (man)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: A54442; S55788; A36317; S55787  
R:Boiger, G.; Michael, T.; Martins, T.; St. John, T.; Steiner, B.; Rodgers, L.; R199  
Mol. Cell. Biol. 13, 6558-6571, 1993  
A:Title: A family of human phosphodiesterases homologous to the dunce learning and me  
A:Reference number: A54442; MUID:94019330  
A:Accession: A54442  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-886 <RES>  
A:Cross-references: GB:L20965; NID:9347119; PIDN:AAA03588.1; PID:9347120  
R:Sullivan, M.; Egerton, M.; Shakur, Y.; Marquardsen, A.; Houslay, M.D.  
Cell. Signal. 6, 793-812, 1994  
A:Title: Molecular cloning and expression, in both COS-1 cells and S. cerevisiae, of  
A:Reference number: S55788; MUID:95194817  
A:Accession: S55788  
A:Molecule type: mRNA  
A:Residues: 'MCPFPVTV', 210-735, 'E', 737-886 <SUL>  
A:Cross-references: EMBL:U18087; NID:9604374; PIDN:AAC50458.1; PID:g604375  
A:Note: 736-Ala was also found  
R:Livi, G.P.; Kmetz, P.; McHale, M.M.; Cieslinski, L.B.; Sathe, G.M.; Taylor, D.P.; D  
Mol. Cell. Biol. 10, 2678-2686, 1990  
A:Title: Cloning and expression of cDNA for a human low-K-m, rolipram-sensitive cycli  
A:Reference number: A36317; MUID:90258854  
A:Accession: A36317  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 'MCPFPVTV', 210-516, 'Y', 518-722, 'R', 724-726, 'R', 728-735, 'E', 737-788, 'E', 7  
A:Cross-references: GB:M37744  
C:Genetics:  
A:Gene: GDB:PDE4A; DPDE2  
A:Cross-references: GDB:138776; OMIM:600126  
A:Map position: 19p13.1-19q12  
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase homology  
C:Keywords: alternative splicing; cAMP binding; phosphoric diester hydrolase  
F:432-660/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 10.5%; Score 111.5; DB 2; Length 886;  
Best Local Similarity 27.5%; Pred. No. 0.27;  
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;

Qy 6 SDVSSSECDREGRLQPAERPPQL-----RCAPTSLQTEPO-----GNPEGNHGG----- 50

Db 49 SUSAERAEERQPHRIERADAMDTSDRCLRTTRMSWFSFGTGTGSGAGGSSRRF 108

Qy 51 EGDSCPHGSP-QGPLAPPASPGPFATRSPLIFMRSSLLSRSSSGYFSFDTDRSPAPMS 109

Query Match	9.9%	Score 105.5;	DB 2;	Length 574;
Best Local Similarity	26.8%	Pred. No. 0.51;		
Matches 41;	Conservative	17;	Mismatches 66;	Indels 29;
				Gaps 6;

RESULT 5  
T42100  
serine/threonine protein kinase (EC 2.7.1.1.) - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
C:Accession: T42100  
R:Bakal, C.J.; Davies, J.E.  
submitted to the EMBL Data Library, September 1998  
A:Description: Cloning, nucleotide sequence and expression of a serine/threonine prot  
A:Reference number: Z22067  
A:Accession: T42100  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-556 <BAK>  
A:Cross-references: EMBL:AF094711; PTDN:AAC64406.1  
C:Genetics:

C;superfamily: synechocystis sp. protein kinase, 58K; protein kinase homology  
C;Keywords: phosphotransferase; serine/threonine-specific protein kinase



R;Mamalaki, A.; Remoundos, M.; Tzartos, S.  
submitted to the EMBL Data Library, May 1995  
A;Description: Molecular cloning of human neuronal nicotinic acetylcholine receptor 4-11  
A;Reference number: S55471  
A;Accession: S55471  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 26-627 <NAM>  
A;Cross-references: EMBL:X87629; NID:g854158; PIDN:CAA60959.1; PID:g854159  
C;Genetics:  
A;Gene: GDB:CHRNA4  
A;Cross-references: GDB:128169; OMIM:118504  
A;Map position: 20q13.2-20q13.3  
C;Superfamily: acetylcholine receptor  
C;Keywords: ion channel; neurotransmitter receptor; postsynaptic membrane; transmembrane

Query Match 9.4%; Score 100; DB 2; Length 627;  
Best Local Similarity 26.4%; Pred. No. 1.5;  
Matches 33; Conservative 15; Mismatches 43; Indels 34; Gaps 5;

QY 1 MAKOPDSVSECDREGLOQAPERPQALRPGAPTSLQTEPOGNGEHNHGSGDSCPHGSP 60  
: l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l  
Db 357 LMKRPSVVKDCNRLLIESMHKMASAPFWP-----EPEGEPPATSGTOSLHPPSPSF 408  
: l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l

QY 61 QGPLAPPASPGPFATRSPLFLIFMRSSLLSRSSGYGFSDTRSP--APMSCDKSTOTPS 118  
: l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l  
Db 409 CVPLDVPAPGCP-SCKSP-----SDQLPPQPPLEAKRASHPS 445  
: l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l l:l:l

QY 119 P-PCQ 122  
: l:l:l  
Db 446 GPCR 450  
: l:l:l

RESULT 10  
S41067  
collagen alpha 1(III) chain - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 13-Jan-1995 #sequence.revision 13-Jan-1995 #text\_change 13-Aug-1999  
C;Accession: S41067; A29905; S31924  
R;Glumoff, V.; Maekelae, J.K.; Vuorio, E.  
Biochim. Biophys. Acta 1217, 41-48, 1994  
A;Title: Cloning of cDNA for rat pro alpha-1(III) collagen mRNA. Different expression pattern  
A;Reference number: S41067; MUID:94114571  
A;Accession: S41067  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-636 <GLU>  
A;Cross-references: EMBL:X70369; NID:g57915; PIDN:CAA49832.1; PID:g57916  
R;Frankel, F.R.; Hsu, C.Y.J.; Myers, J.C.; Lin, E.; Lyttle, C.R.; Komm, B.; Mohn, K.  
DNA 7, 347-354, 1988  
A;Title: Regulation of alpha-2 (I), alpha-1 (III), and alpha-2(V) collagen mRNAs by estradiol  
A;Reference number: A29905; MUID:88296083  
A;Accession: A29905  
A;Molecule type: mRNA  
A;Residues: 308-482 <FRA>  
A;Cross-references: GB:M21354; NID:g203500; PIDN:AAA40942.1; PID:g203501  
R;Glumoff, V.; Maekelae, J.K.; Vuorio, E.  
submitted to the EMBL Data Library, February 1993  
A;Reference number: S31924  
A;Accession: S31924  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 2-636 <GL2>  
A;Cross-references: EMBL:X70369  
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix  
F;408-636/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 9.3%; Score 99; DB 2; Length 636;  
Best Local Similarity 35.4%; Pred. No. 1.9;  
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;



submitted to the EMBL Data Library, February 1989

A:Reference number: S05272  
A:Accession: S05272  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1240, 'V', 1242-1466 <RCS>  
A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1.; PID:g30058  
R:Ala-Kokko, L.; Kontusaari, S.; Baldwin, C.T.; Kulvaneniemi, H.; Prockop, D.J.  
Biochem. J. 260, 509-516, 1989  
A:Title: Structure of cDNA clones coding for the entire prepro-alpha(III) chain of human  
enkephalins.  
A:Reference number: S04642; MUID:89350838  
A:Accession: S04642  
A:Molecule type: mRNA  
A:Residues: 1-1196 <ALA>  
A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1.; PID:g30058  
A:Note: the complete sequence is not shown  
R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.  
Gene 78, 255-265, 1989  
A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene  
A:Reference number: P80011; MUID:89378752  
A:Accession: P80011  
A:Molecule type: DNA  
A:Residues: 1-176 <BEN>  
A:Cross-references: GB:M26939; NID:g180813; PIDN:AAA52040.1.; PID:g180814  
R:Toman, P.D.; Ricca, G.A.; de Crombrughe, B.  
Nucleic Acids Res. 16, 7201, 1988  
A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pre  
A:Reference number: S01726; MUID:88303360  
A:Accession: S01726  
A:Molecule type: mRNA  
A:Residues: 1-170 <TON>  
A:Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1.; PID:g30061  
A:Note: the authors translated the codon CAG for residue 154 as H1S  
R:Janeczko, R.A.; Ramirez, F.  
Nucleic Acids Res. 17, 6742, 1989  
A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen.  
A:Reference number: S04887; MUID:89386015  
A:Accession: S04887  
A:Molecule type: mRNA  
A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634,  
A:Cross-references: EMBL:X15332; NID:g29545; PIDN:CAA33387.1.; PID:g9930045  
A:Note: the authors' translation of residues 905-932 is inconsistent with the nucleotide  
R:Seyer, J.M.; Kang, A.H.  
Biochemistry 16, 1158-1164, 1977  
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptide  
A:Reference number: A90399; MUID:77134724  
A:Accession: A90399  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>  
A:Experimental source: liver  
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact  
R:Seyer, J.M.  
submitted to the Atlas, December 1977  
A:Reference number: A94562  
A:Accession: A94562  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>  
A:Experimental source: liver  
A:Note: author submitted corrections to A90399  
R:Milwicz, D.M.; Wit, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.  
Am. J. Hum. Genet. 53, 62-70, 1993  
A:Title: Parental somatic and germ-line mosaicism for a multiexon deletion with unusual  
splicing.  
A:Reference number: I51868; MUID:93304430  
A:Accession: I51868  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 186-194 <MIL>  
A:Cross-references: GB:S62925; NID:g386425; PIDN:AAD13937.1.; PID:g4261637  
R:Chiodo, A.A.; Sillence, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3

A:Reference number: S59511; MUID:96067614  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 302-423 <CHI>  
A:Cross-references: GB:S79877; NID:g1195576; PIDN:AAB35615.1.; PID:g1195577  
R:Seyer, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978  
A:Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBR  
A:Reference number: A90414; MUID:79000343  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5256-5259, 1991  
A:Title: G to T transition at position +5 of a splice donor site causes skipping of  
A:Reference number: I55349; MUID:91161621  
A:Accession: I55349  
A:Status: translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 537-605 <LEE>  
A:Cross-references: GB:M59312; NID:g180815; PIDN:AAA52041.1.; PID:g180816  
R:Seyer, J.M.; Mainardi, C.; Kang, A.H.  
Biochemistry 19, 1583-1589, 1980  
A:Title: Covalent structure of collagen: amino acid sequence of alpha (III)-CB5 from  
A:Reference number: A90438; MUID:80198282  
A:Accession: A90438  
A:Molecule type: protein  
A:Residues: 728-895, 'A', 897-964 <SEY4>  
A:Experimental source: liver  
R:Cole, W.G.; Chiodo, A.A.; Lamande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; C  
J. Biol. Chem. 265, 17070-17077, 1990  
A:Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping  
A:Reference number: A38303; MUID:91009133  
A:Accession: A38303  
A:Molecule type: mRNA  
A:Residues: 861-1015 <COL>  
A:Cross-references: GB:J05617; GB:M55603; GB:M59227; NID:g180878; PIDN:AAB59383.1.; PI  
A:Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos  
R:ManKoo, B.S.; Dalgleish, R.  
Nucleic Acids Res. 16, 2337, 1988  
A:Title: Human pro alpha(I(III)) collagen: cDNA sequence for the 3' end.  
A:Reference number: S02119; MUID:88189827  
A:Accession: S02119  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>  
A:Cross-references: EMBL:X06700; NID:g30053; PIDN:CAA29886.1.; PID:g30054  
R:Seyer, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A:Title: Covalent structure of collagen: amino acid sequence of alpha (III)-CB9 from  
A:Reference number: A90446; MUID:81208139  
A:Accession: A90446  
A:Molecule type: protein  
A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 11  
A:Experimental source: liver  
R:Loidl, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.;  
Nucleic Acids Res. 12, 9383-9394, 1984  
A:Title: Molecular cloning and carboxyl-propeptide analysis of human type III procoll  
A:Reference number: A93551; MUID:85087944  
A:Accession: A93551  
A:Molecule type: mRNA  
A:Residues: 1065-1155, 'P', 1157-1466 <LOI>  
A:Cross-references: EMBL:X01655; EMBL:X01742; NID:g29584; PIDN:CAA25821.1  
R:Wiskuln, M.; Dalgleish, R.; Kluve-Beckerman, B.; Rennard, S.I.; Tolstoshev, P.; Br  
Biochemistry 25, 1408-1413, 1986  
A:Title: Human type III collagen gene expression is coordinately modulated with the t  
A:Reference number: I52393; MUID:86187804  
A:Accession: I52393  
A:Status: translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1161-1200 <MIS>  
A:Cross-references: GB:M13146; NID:g180415; PIDN:AAA52003.1.; PID:g180416



C:Accession: T10811  
 R:Irrie, M.; Hata, Y.; Takai, Y.  
 A:Submitted to the EMBL Data Library, April 1996  
 A:Description: Cloning of new isoforms of PSD-95/SAP90 related genes.  
 A:Reference number: Z17166  
 A:Accession: T10811  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-852 <IRI>  
 A:Cross-references: EMBL:U53368; NID:g1517939; PID:g1517940  
 C:Genetics:  
 A:Note: PSD-95/SAP90-related gene 2 (chapsyn isoform2)  
 C:Superfamily: discs-large tumor suppressor; GLGF domain homology; guanylate kinase homology  
 F:198-216/Domain: GLGF domain homology <GLG>  
 F:543-601/Domain: SH3 homology <SH3>  
 F:663-840/Domain: guanylate kinase homology <GKI>

Query Match 9.1%; Score 97; DB 2; Length 852;  
 Best Local Similarity 23.8%; Pred. No. 3.7;  
 Matches 44; Conservative 21; Mismatches 58; Indels 62; Gaps 9;

QY 28 LRPGAPTSL-QTEPOGNPGNHGGEGDCSPH--GSPQGPL-----APPASPGPEA----- 74  
 DB 274 LKVGRTTIYWDPIGPPDITHSYSPPMENHLLSGNNGTLEYKTSLLPISGRISPIPKH 333  
 QY 75 -----TRSPLEIFMRRSSLLSRSSSYFSFTDRSPAPMSCDKS--TQTPSPPCQAFN 125  
 DB 334 MLVEDEYTRPPEPVYSTVKNKCDKPAS-----PRHYSFVECDKSFLLSTYP----- 380  
 QY 126 HYL-----SAMASMRQAPADMRPEIWTIAQLRIGDEFNAYYARRVFLNNYQAAEDHP 179  
 DB 381 HYHGLLPDSDMTSHSQHSTATROPSPVTLQRAI-----SLEGE 419  
 QY 180 RMVIL 184  
 DB 420 RKWL 424

Search completed: August 16, 2002, 10:01:47  
 Job time: 140 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 10:01:22 ; Search time 11.99 Seconds  
(without alignments)  
639.405 Million cell updates/sec

Title: US-09-508-832-10  
Perfect score: 1065  
Sequence: 1 MAKOPSDVSECDREGRLQ.....PRMVLRLRYIVLVRMH 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Length	DB	ID	Description
1	1065	100.0	198	1	BIM_HUMAN	O43521 homo sapien
2	908	85.3	196	1	BIM_MOUSE	O54918 mus musculus
3	893	83.8	196	1	BIM_RAT	O88498 rattus norv
4	111.5	10.5	886	1	CNA4_HUMAN	P27815 homo sapien
5	106.5	10.0	620	1	EXON_HSV2	P06489 herpes simp
6	102.5	9.6	1395	1	CUT1_MOUSE	P53564 mus musculus
7	100	9.4	627	1	ACHA_HUMAN	P43681 homo sapien
8	99	9.3	636	1	CA13_RAT	P13941 rattus norv
9	99	9.3	1464	1	CA13_MOUSE	P08121 mus musculus
10	99	9.3	1466	1	CA13_HUMAN	P02461 homo sapien
11	98	9.2	105	1	COLL_HSV7	P25050 herpesvirus
12	97.5	9.2	1505	1	CUT1_HUMAN	P39880 homo sapien
13	97	9.1	852	1	DLG2_RAT	O63622 rattus norv
14	96	9.0	870	1	DLG2_HUMAN	Q15700 homo sapien
15	96	9.0	1527	1	CA1H_MOUSE	P39061 mus musculus
16	95.5	9.0	210	1	CAC2_HAECO	P16252 haemonchus
17	95	8.9	298	1	CC34_CAEEL	P34687 caenorhabdi
18	95	8.9	844	1	CNA4_RAT	P54748 rattus norv
19	95	8.9	1516	1	CA1H_HUMAN	P39060 homo sapien
20	94.5	8.9	573	1	SECD_MYCTU	Q50634 mycobacteri
21	94.5	8.9	628	1	V7OK_TYMW	P20131 turnip yell
22	94.5	8.9	1183	1	DRPL_RAT	P54258 rattus norv
23	94.5	8.9	1185	1	DRPL_HUMAN	P54259 homo sapien
24	94.5	8.9	1234	1	B3A2_RAT	P23347 rattus norv
25	94	8.8	580	1	SYN3_HUMAN	O14994 homo sapien
26	94	8.8	1460	1	CA11_CANFA	Q9X5J7 canis fami
27	93.5	8.8	902	1	NFC4_HUMAN	Q14934 homo sapien
28	92.5	8.7	245	1	C1QA_HUMAN	P02745 homo sapien
29	92.5	8.7	408	1	AL_DROME	Q06453 drosophila
30	92.5	8.7	1446	1	IE18_PVKKA	P33479 pseudorabie
31	92.5	8.7	1461	1	IE18_PRVIF	P11675 pseudorabie
32	91.5	8.6	234	1	PRPM_HUMAN	P10161 homo sapien
33	91.5	8.6	276	1	PRPL_HUMAN	P10162 homo sapien

## ALIGNMENTS

## RESULT 1

ID	BIM_HUMAN	STANDARD;	PRT;	198 AA.
AC	O43521; O43522;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Bcl2-like protein 11 (Bcl2 interacting mediator of cell death).			
GN	BCL2L11 OR BIM.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid:9606;			
[1]				
RP	SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE SPLICING.			
RC	TISSUE=Peripheral blood, and Spleen;			
RX	MEDLINE=98094360; PubMed=9430630;			
RA	O'Connor L., Strasser A., O'Reilly L.A., Hausmann G., Adams J.M.,			
RA	Cory S., Huang D.C.S.;			
RT	*Bim: a novel member of the Bcl-2 family that promotes apoptosis.*;			
RL	EMBO J. 17:384-395(1998).			
CC	-!- FUNCTION: INDUCES APOPTOSIS. ISOFORM BIML IS MORE POTENT THAN			
CC	ISOFORM BIMEL.			
CC	-!- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2			
CC	PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES			
CC	NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,			
CC	BAX OR BAK (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOSOLASMIC MEMBRANES			
CC	(BY SIMILARITY).			
CC	-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BIMEL (SHOWN HERE) AND			
CC	BIML; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND			
CC	CYTOTOXICITY.			
CC	-!- SIMILARITY: CONTAINS 1 BCL-2 HOMOLGY DOMAIN 3 (BH3).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AF032457; AAC39593.1; -.			
DR	EMBL; AF032458; AAC39594.1; -.			
DR	MIM; 603827; -.			
DR	InterPro; IPR000712; Bcl_2.			
DR	PROSITE; PS01259; BH3; FALSE_NEG.			
KW	Apoptosis; Alternative splicing; Membrane.			
FT	DOMAIN 148 162 BH3.			
FT	VARSPLIC 42 101 MISSING (IN ISOFORM BIML).			
SQ	SEQUENCE 198 AA; 22171 MW; D75735B469CA6997 CRC64;			

Query Match 100.0%; Score 1065; DB 1; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1e-68;

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKOPSDVSSECDREGQLQPAERPPQLRPGAPTSLOTEPOGNEGNHGGEGDSCPHGSP 60  
 DB 1 MAKOPSDVSSECDREGQLQPAERPPQLRPGAPTSLOTEPOGNEGNHGGEGDSCPHGSP 60

QY 61 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120  
 DB 61 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120

QY 121 COAFNHVLSAMASMRQAEPADMRPEIWAQELRRIGDEFNAYYARRVFLNNYQAAEDHPR 180  
 DB 121 COAFNHVLSAMASMRQAEPADMRPEIWAQELRRIGDEFNAYYARRVFLNNYQAAEDHPR 180

QY 181 MVILRLRYIVRLVWRMH 198  
 DB 181 MVILRLRYIVRLVWRMH 198

RESULT 2

BIM\_MOUSE  
 ID BIM\_MOUSE STANDARD; PRT; 196 AA.  
 AC 054918; 054920;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE BCL2-like protein 11 (BCL2 interacting mediator of cell death).  
 GN BCL2L11 OR BIM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE  
 SPECIFICITY, AND ALTERNATIVE SPLICING.  
 RX MEDLINE=98094360; PubMed=9430630;  
 RA O'Connor L., Strasser A., O'Reilly L.A., Hausmann G., Adams J.M.,  
 RA Cory S., Huang D.C.S.;  
 RL "Bim: a novel member of the Bcl-2 family that promotes apoptosis.";  
 EMBO J. 17:384-395(1998).  
 CC -1- FUNCTION: INDUCES APOPTOSIS. THE ISOFORMS VARY IN CYTOTOXICITY  
 WITH ISOFORM BIMs BEING THE MOST POTENT AND ISOFORM BIMEL BEING  
 THE LEAST POTENT.  
 CC -1- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2  
 PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES  
 NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,  
 BAX OR BAK (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOSOLIC MEMBRANES.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; BIMEL (SHOWN HERE), BIML AND  
 BIMs; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF B-AND T-LYMPHOID CELL  
 LINES.  
 CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND  
 CYTOTOXICITY.  
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLGY DOMAIN 3 (BH3).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AF032459; AAC40029.1; -  
 DR EMBL: AF032460; AAC40030.1; -  
 DR EMBL: AF032461; AAC40031.1; -  
 DR MGD; MGI:1197519; Bcl2l11.  
 DR InterPro; IPR000712; Bcl\_2.  
 DR PROSITE; PS01259; BH3; FALSE\_NEG.  
 KW Apoptosis; Alternative splicing; Membrane.  
 FT DOMAIN 146 160 BH3  
 FT VARSPPLIC 42 97 MISSING (IN ISOFORM BIML).

FT VARSPPLIC 42 127 MISSING (IN ISOFORM BIMs).  
 SQ SEQUENCE 196 AA; 22066 MW; 531C176E5FIAC9AA CRC64;

Query Match 85.3%; Score 908; DB 1; Length 196;  
 Best Local Similarity 86.5%; Pred. No. 1.1e-57;  
 Matches 173; Conservative 10; Mismatches 11; Indels 6; Gaps 2;

QY 1 MAKOPSDVSSECDREGQLQPAERPPQLRPGAPTSLOTEPOGNEGNHGGEGDSCPHGSP 60  
 DB 1 MAKOPSDVSSECDREGQLQPAERPPQLRPGAPTSLOTEPOGNEGNHGGEGDSCPHGSP 56

QY 61 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120  
 DB 57 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 116

QY 121 COAFNHVLSAMASMRQAEPADMRPEIWAQELRRIGDEFNAYYARRVFLNNYQAAEDH 178  
 DB 117 COAFNHVLSAMASMRQAEPADMRPEIWAQELRRIGDEFNAYYARRVFLNNYQAAEDH 176

QY 179 PRVILRLRYIVRLVWRMH 198  
 DB 177 PQWILQLLRIFELVWRMH 196

RESULT 3

BIM\_RAT  
 ID BIM\_RAT STANDARD; PRT; 196 AA.  
 AC 088498; 09WU18; 088497;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE BCL2-like protein 11 (BCL2 interacting mediator of cell death)  
 DE (Bcl-2 related ovarian death protein).  
 GN BCL2L11 OR BIM OR BOD.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND TISSUE SPECIFICITY  
 (ISOFORMS BOD-L; BOD-M AND BOD-S).  
 RC TISSUE-Ovary;  
 RX MEDLINE=98400436; PubMed=9731710;  
 RA Hsu S.Y., Lin P., Hsueh A.J.W.;  
 RT "BOD (Bcl-2-related ovarian death gene) is an ovarian BH3 domain-  
 containing proapoptotic Bcl-2 protein capable of dimerization with  
 diverse antiapoptotic Bcl-2 members.";  
 RL Mol. Endocrinol. 12:1432-1440(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM BIML).  
 RA Chen D., Simon R.P., Chen J.;  
 RT "Cloning of rat bimEL and bimL, and their differential expression in  
 ischemia and normal rat brain.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INDUCES APOPTOSIS.  
 CC -1- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2  
 PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES  
 NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,  
 BAX OR BAK.  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOSOLIC MEMBRANES  
 (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BOD-L (SHOWN HERE) AND BOD-S;  
 ARE PRODUCED BY THE USE OF ALTERNATIVE INITIATION SITES. TWO  
 FURTHER ISOFORMS; BIML AND BOD-M; ARE PRODUCED BY ALTERNATIVE  
 SPLICING OF BOD-L.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND  
 CYTOTOXICITY.  
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLGY DOMAIN 3 (BH3).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AF065433; AAC23595.1; -  
DR EMBL: AF065431; AAC23593.1; -  
DR EMBL: AF065432; AAC23594.1; -  
DR EMBL: AF136927; AAD26594.1; -  
DR InterPro: IPR000712; BCL\_2  
DR PROSITE: PS01259; BH3; FALSE\_NEG.  
KW Apoptosis; Alternative splicing; Membrane; Alternative initiation.  
FT CHAIN 1 196 BCL2-LIKE PROTEIN 11, ISOFORM BOD-L.  
FT CHAIN 104 196 BCL2-LIKE PROTEIN 11, ISOFORM BOD-S.  
FT INIT\_MET 104 104 FOR ISOFORM BOD-S.  
FT DOMAIN 146 160 BH3.  
FT VARSPPLIC 42 97 MISSING (IN ISOFORM B1ML).  
FT VARSPPLIC 42 137 MISSING (IN ISOFORM BOD-M).  
FT CONFLICT 136 136 E -> D (IN REF. 1; AAC23594).  
SQ SEQUENCE 196 AA; 22055 MW; B4D2146F9C0B37A0 CRC64;

Query Match 83.8%; Score 893; DB 1; Length 196;  
Best Local Similarity 85.0%; Pred. No. 1.2e-56;  
Matches 170; Conservative 11; Mismatches 13; Indels 6; Gaps 2;

QY 1 MAKOPSDVSECDREGCLOQAPRRPQLRGAPTSLQTEPQGNHGGDCSPHGSP 60  
DB 1 MAKOPSDVSECDREGCLOQAPRRPQLRGAPTSLQTESQGNPD---GEGDRCSPHGSP 56  
QY 61 QGCLAPPASGPGPATRSPFLIFMRRSLLSRSSGYSFSDTDRSPAPMSCDKSTQTPSP 120  
DB 57 QGCLAPPASGPGPATRSPFLIFVRRSLLSRSSGYSFSDTDRSPAPMSCDKSTQTPSP 116  
QY 121 QCAFNYHLSMASMROA--EPADMRREITWQELRIGRIGDEFNAYYARRVFLNNYQAEHD 178  
DB 117 QCAFNYHLSMASIRSQSEPEDLRPEIRTAQELRIGRIGDEFNFTYTRRAFANDYREAEDH 176  
QY 179 PRMVIILRLRYIVRLVWRMH 198  
DB 177 PQWVILQLLRFIFRLVWRH 196

RESULT 4  
CN4A\_HUMAN STANDARD: PRT: 886 AA.  
ID CN4A\_HUMAN Q16691; Q16255; O75522; O76092;  
AC P27815; Q16691; Q16255; O75522; O76092;  
DT 01-OCT-1992 (Rel. 23, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CAMP-dependent 3',5'-cyclic phosphodiesterase 4A (EC 3.1.4.17)  
DE (DPDE2) (PDE46).  
GN PDE4A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94019330; PubMed=8413254;  
RA Bolger G., Michaeli T., Martins T., St John T., Steiner B.,  
RA Rodgers L., Riggs M., Wigler M., Ferguson K.;  
RT "A family of human phosphodiesterases homologous to the dunce  
RT learning and memory gene product of Drosophila melanogaster are  
RT potential targets for antidepressant drugs.";  
RL Mol. Cell. Biol. 13:6558-6571(1993).  
RN [2]  
RP SEQUENCE FROM N.A. (PDE4A RD1).  
RC TISSUE=Brain;  
RX MEDLINE=98343959; PubMed=9677330;  
RA Sullivan M., Rana G., Begg F., Gordon L., Olsen A.S., Houslay M.D.;  
RT "Identification and characterization of the human homologue of the

RT short PDE4A CAMP-specific phosphodiesterase RD1 (PDE4A1) by analysis  
RT of the human HSPD4A gene locus located at chromosome 19p13.2.";  
RL Biochem. J. 333:693-703(1998).  
RN [3]  
RP SEQUENCE OF 112-886 FROM N.A.  
RC TISSUE=Monocytes;  
RX MEDLINE=90258854; PubMed=2160582;  
RA Livi G.P., Kmetz P., McHale M.M., Cieslinski L.B., Sathe G.M.,  
RA Taylor D.P., Davis R.L., Torphy T.J., Balcarek J.M.;  
RT "Cloning and expression of cDNA for a human low-Km, rolipram-sensitive  
RT cyclic AMP phosphodiesterase.";  
RL Mol. Cell. Biol. 10:2678-2686(1990).  
RN [4]  
RP REVISIONS TO REF.3.  
RA McLaughlin M.M.;  
RL Submitted (JAN-1991) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (PDE4A7/A6 AND PDE4A8/2EL).  
RX MEDLINE=95290008; PubMed=7772058;  
RA Horton Y.M., Sullivan M., Houslay M.D.;  
RT "Molecular cloning of a novel splice variant of human type IVA  
RT (PDE-IVA) cyclic AMP phosphodiesterase and localization of the gene  
RT to the p13.2-q12 region of human chromosome 19.";  
RL Biochem. J. 308:683-691(1995).  
RN [6]  
RP SEQUENCE FROM N.A. (PDE4A7/A6).  
RX MEDLINE=95194817; PubMed=7888306;  
RA Sullivan M., Egerton M., Shakur Y., Marquardsen A., Houslay M.D.;  
RT "Molecular cloning and expression, in both COS-1 cells and S.  
RT cerevisiae, of a human cytosolic type-IVA, cyclic AMP specific  
RT phosphodiesterase (hPDE-IVA-h6.1).";  
RL Cell. Signal. 6:793-812(1994).  
CC -1- CATALYTIC ACTIVITY: ADENOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O -  
CC ADENOSINE 5'-PHOSPHATE.  
CC -1- ENZYME REGULATION: INHIBITED BY ROLIPRAM.  
CC -1- PATHWAY: CYCLIC NUCLEOTIDE METABOLISM.  
CC -1- SUBCELLULAR LOCATION: THE PDE4A1 RD1 ISOFORM HAS PROPENSITY FOR  
CC ASSOCIATION WITH MEMBRANES.  
CC -1- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF THE PROTEIN ARE PRODUCED BY  
CC AND PDE4A1/RD1 A SHORT ISOFORM. ISOFORM 2EL PROBABLY REPRESENTS A  
CC NON-FUNCTIONAL SPICE VARIANT. THE PRESENCE OF DIFFERENT N-  
CC TERMINAL REGIONS HAS BEEN LINKED WITH DISTINCT FUNCTIONAL ROLES.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
CC FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: L20965; AAA03588.1; -  
DR EMBL: AF069491; AAC35012.1; -  
DR EMBL: AF069487; AAC35012.1; JOINED.  
DR EMBL: AF069489; AAC35012.1; JOINED.  
DR EMBL: AF069490; AAC35012.1; JOINED.  
DR EMBL: AF069491; AAC35013.1; -  
DR EMBL: AF069489; AAC35013.1; JOINED.  
DR EMBL: AF069490; AAC35013.1; JOINED.  
DR EMBL: AF069491; AAC35014.1; -  
DR EMBL: AF069489; AAC35014.1; JOINED.  
DR EMBL: AF069490; AAC35014.1; JOINED.  
DR EMBL: M37744; AAA69697.1; -  
DR EMBL: U18087; AAC50458.1; -  
DR EMBL: U18088; AAC50458.1; -  
DR EMBL: S75213; AAB33798.1; -  
DR EMBL: AF069491; AAC35015.1; -  
DR EMBL: AF069488; AAC35015.1; JOINED.  
DR EMBL: AF069489; AAC35015.1; JOINED.  
DR EMBL: AF069490; AAC35015.1; JOINED.





[2]  
RP SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).  
RC STRAIN-A/J. AND BALB/C: TISSUE=Brain;  
RX MEDLINE=94244481; PubMed=7910552;  
RA Valarche I.; Tissier-Seca J.P.; Hirsch M.R.; Martinez S.; Goridis C.;  
RN Brunet J.F.;  
RT "The mouse homeodomain protein Phox2 regulates Ncam promoter activity  
in concert with Cux/CDP and is a putative determinant of  
neurotransmitter phenotype.";  
RL Development 119:881-896(1993).  
RN [3]  
RP SEQUENCE OF 642-1395 FROM N.A.  
RX MEDLINE=96437626; PubMed=8840273;  
RA Den Heuvel G.B.; Bodmer R.; McConnell K.R.; Nagami G.T.; Igarashi P.;  
RT "Expression of a cut-related homeobox gene in developing and  
polycystic mouse kidney.";  
RL Kidney Int. 50:453-461(1996).  
RN [4]  
RP SEQUENCE OF 936-1395 FROM N.A.  
RX TISSUE=Testis;  
RA Quaglin S.E.; Igarashi P.;  
RT "A unique variant of a homeobox gene related to Drosophila cut is  
expressed in mouse testis.";  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A  
REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY  
PREVENTING BINDING OF POSITIVELY-ACTING CCAAT FACTORS TO  
PROMOTERS (BY SIMILARITY). COMPONENT OF NF-MUNR REPRESSOR; BINDS  
TO THE MARS (5' AND 3') OF THE IMMUGLOBULIN HEAVY CHAIN  
ENHANCER.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.  
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEOBOX PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF004225; AAD12485.1; -.  
DR EMBL; X75013; CAA52922.1; -.  
DR EMBL; U46683; AAC52775.1; -.  
DR EMBL; U46684; AAB41146.1; -.  
DR HSSP; P10037; 1AU7.  
DR MGD; MGI:88568; Cut11.  
DR InterPro; IPR003350; CUT.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF02376; CUT; 3.  
DR Pfam; PF00046; homeobox; 2.  
DR SMART; SM00389; HOX; 1.  
DR PROSITE; PS00027; HOMEOBOX\_1; 1.  
DR PROSITE; PS50071; HOMEOBOX\_2; 1.  
KW Transcription regulation; Homeobox; DNA-binding;  
KW Developmental protein; Nuclear protein; Repeat; Repressor;  
KW Coiled coil; Alternative splicing.  
FT NON\_TER 1 1  
FT DOMAIN 1 243 COILED COIL (POTENTIAL).  
FT DNA\_BIND 420 507 CUT 1.  
FT DOMAIN 547 603 COILED COIL (POTENTIAL).  
FT DNA\_BIND 809 896 CUT 2.  
FT DNA\_BIND 992 1079 CUT 3.  
FT DNA\_BIND 1119 1178 HOMEOBOX.  
FT VARSPPLIC 287 388 MISSING (IN ISOFORM 2).  
FT CONFLICT 1360 1360 G -> A (IN REF. 2).  
FT CONFLICT 1385 1385 P -> L (IN REF. 1).  
SQ SEQUENCE 1395 AA; 151802 MW; D062CC227D7A163E CRC64;

Query Match 9.6%; Score 102.5; DB 1; Length 1395;  
Best Local Similarity 23.7%; Pred. No. 3.5;  
Matches 54; Conservative 25; Mismatches 76; Indels 73; Gaps 9;  
QY 6 SVSSECDREGHQLOPAERPPOLRPGAPVSLQT---EPQGNPEGNHGGEDSCPHGSPG 62  
Db 282 SLSGSARSKRGDQESRRPGPLPASPPLPRTNGTGEQVSTNGTH-----HFSPAG 333  
QY 63 -----PLAPPASPGPFATRSPLFIEMRS-----SLLSRSSSGVFSFD 100  
Db 334 LSQDFSSWLASPSLPLASTGRFALNLLQRLQMOSFYSKAMQEAAGSTIFSTGPISTN 393  
QY 101 TDRSPAPM---SCDKSTQTPSP-----COAFN 125  
Db 394 SISSPSLQOSPDVNCMAPSPQSESAGSISGEEIDTAIEARQVKEQLIKHNGIRIFG 453  
QY 126 HYLASAMASMRQAEPADMRPEIWIQAELRIGDEFNAYARRVFLNNYQ 173  
Db 454 HYVLGLSQGSVSEIL-ARPKPNKLTVR--GKE--PFHKMKOFLSDEQ 496  
RESULT 7  
ACH4\_HUMAN  
ID AC44\_HUMAN STANDARD; PRT; 627 AA.  
AC P43681;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Neuronal acetylcholine receptor protein, alpha-4 chain precursor.  
GN CHRNA4 OR NACRA4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=95237608; PubMed=7721089;  
RA Monteggia L.M.; Gopalakrishnan M.; Touma E.; Idler K.B.; Nash N.;  
RA Arneric S.P.; Sullivan J.P.; Giordano T.;  
RT "Cloning and transient expression of genes encoding the human alpha-4  
and beta-2 neuronal nicotinic acetylcholine receptor (nAChR)  
subunits.";  
RT Gene 155:189-193(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96430009; PubMed=88333159;  
RA Steinlein O.K.; Weiland S.; Stood J.; Propping P.;  
RT "Exon-intron structure of the human neuronal nicotinic acetylcholine  
receptor alpha 4 subunit (CHRNA4).";  
RL Genomics 32:289-294(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Hippocampus;  
RX MEDLINE=97062879; PubMed=8906617;  
RA Elliott K.J.; Ellis S.B.; Berckhan K.J.; Urrutia A.;  
RA Chavez-Noriega L.E.; Johnson E.C.; Velicelebi G.; Harpold M.M.;  
RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta  
2-beta 4 nicotinic acetylcholine receptor subunits and functional  
expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and  
beta 4 subunits.";  
RT J. Mol. Neurosci. 7:217-228(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97162233; PubMed=9009220;  
RA Groot Kormelink P.J.; Luyten W.H.M.L.;  
RT "Cloning and sequence of full-length cDNAs encoding the human neuronal  
nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and  
expression of seven nAChR subunits in the human neuroblastoma cell  
line SH-SY5Y and/or IMR-32.";  
RT FEBS Lett. 400:309-314(1997).  
RN [5]  
RP SEQUENCE FROM N.A.

[illegible]

```

RA Komm B., Mohn K.;
RT "Regulation of alpha 2(I), alpha 1(III), and alpha 2(V) collagen
RL mRNAs by estradiol in the immature rat uterus.";
CC DNA 7:347-354(1988).
CC -!- FUNCTION: COLLAGEN.
CC -!- ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70369; CAA49832.1; -.
DR EMBL; AJ005395; CAA06510.1; -.
DR EMBL; M21354; AAA40942.1; -.
DR PIR; A29905; A29905.
DR PIR; S41067; S41067.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib.collagen_C.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 6.
DR ProDom; PD002078; Fib.collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR PROSITE; PS01208; WFPC; PARTIAL.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Collagen; Glycoprotein.
FT NON_TER 1
FT CHAIN <1 375 COLLAGEN ALPHA 1(III) CHAIN.
FT PROPEP 376 636 CARBOXYL-TERMINAL PROPEPTIDE.
FT DOMAIN <1 368 TRIPLE-HELICAL REGION.
FT DOMAIN 369 636 NONHELICAL REGION (C-TERMINAL).
FT DISULFID 368 368 INTERCHAIN (BY SIMILARITY).
FT DISULFID 369 369 INTERCHAIN (BY SIMILARITY).
FT CONFLICT 340 420 N -> D (IN REF. 2).
FT CONFLICT 429 429 A -> G (IN REF. 2).
SQ SEQUENCE 636 AA; 62332 MW; 61AA48159F01D01EE CRC64;

Query Match 9.3%; Score 99; DB 1; Length 636;
Best Local Similarity 35.4%; Pred. No. 2.9;
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;

QY 5 PSDVSSECDREGROLQPAER---PPQLRPGAP-TSLQTEPQGNP-----EGNHGGEG 52
Db 145 PQGIKESKPGKASNGHGERPGPGGLPGQPGTAGEPDGNPGSDGPGRDGSPGGK 204
QY 53 DSCPFGSPQGPLAP--PASPGP 72
Db 205 DRNGSPGAPGAPGHPGPPG 226

RESULT 9
CAL3_MOUSE STANDARD; PRT: 1464 AA.
AC P08121; O61429; O9CRN7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(III) chain precursor.
GN COL3A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL X DBA; TISSUE=Embryo;
RX MEDLINE=95011609; PubMed=7926795;
RA Toman D., de Crombrughe B.;
RT "The mouse type-III procollagen-encoding gene: genomic cloning and
RL complete DNA sequence.";
RN [2]
RP SEQUENCE OF 1-488 FROM N.A.
RX MEDLINE=88167858; PubMed=3443309;
RA Wood L., Theriault N., Vogel G.;
RT "Complete nucleotide sequence of the N-terminal domains of the murine
RL alpha-1 type-III collagen chain.";
RN [3]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=85131189; PubMed=3972847;
RA Liau G., Mudryj M., de Crombrughe B.;
RT "Identification of the promoter and first exon of the mouse alpha 1
RL (III) collagen gene.";
RN [4]
RP SEQUENCE OF 810-1464 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [5]
RP SEQUENCE OF 1442-1464 FROM N.A.
RC STRAIN=C57BL;
RX MEDLINE=91274355; PubMed=2054384;
RA Metsaeranta M., Toman D., de Crombrughe B., Vuorio E.;
RT "Specific hybridization probes for mouse type I, II, III and IX
RL collagen mRNAs.";
CC -!- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
CC ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC O-LINKED GLYCAN CONSISTS OF GLC-GAL DISACCHARIDE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 WFCC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52046; CAA36279.1; -.

```



RA Miskulin M., Dalgleish R., Kluve-Beckerman B., Rennard S.I.,  
RA Tolstoshev P., Brantly M., Crystal R.G.;  
RT "Human type III collagen gene expression is coordinately modulated  
RT with the type I collagen genes during fibroblast growth.";  
RL Biochemistry 25:1408-1413(1986).  
RN [13]  
RP SEQUENCE OF 1-170 FROM N.A.  
RC TISSUE-Placenta;  
RX MEDLINE=88303360; PubMed=3405773;  
RA Toman D., Ricca G., de Crombrughe B.;  
RT "Nucleotide sequence of a cDNA coding for the amino-terminal region  
RT of human prepro alpha 1(III) collagen.";  
RL Nucleic Acids Res. 16:7201-7201(1988).  
RN [14]  
RP SEQUENCE OF 1-176 FROM N.A.  
RX MEDLINE=89378752; PubMed=2777083;  
RA Benson-Chanda V., Su M.W., Weil D., Chu M.-L., Ramirez F.;  
RT "Cloning and analysis of the 5' portion of the human type-III  
RT procollagen gene (COL3A1).";  
RL Gene 78:255-265(1989).  
RN [15]  
RP REVIEW ON VARIANTS.  
RX MEDLINE=97255959; PubMed=9101290;  
RA Kuivaniemi H., Tromp G., Prockop D.J.;  
RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-  
RT associated collagen (type IX), and network-forming collagen (type X)  
RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";  
RL Hum. Mutat. 9:300-315(1997).  
RN [16]  
RP VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668.  
RX MEDLINE=93293988; PubMed=8514866;  
RA Tromp G., Wu Y., Prockop D.J., Madhathari S.L., Kleinert C.,  
RA Earley J.J., Zhuang J., Noerregaard O., Darling R.C., Abbott W.M.,  
RA Cole C.W., Jaakkola P., Rynnaeren M., Pearce W.H., Yao J.S.T., S.A.,  
RA Majumaa K., Smullens S.N., Gatalica Z., Ferrell R.E., Jimenez S.A.,  
RA Jackson C.E., Michels V.V., Kaye M., Kuivaniemi H.;  
RT "Sequencing of cDNA from 50 unrelated patients reveals that mutations  
RT in the triple-helical domain of type III procollagen are an  
RT infrequent cause of aortic aneurysms.";  
RL J. Clin. Invest. 91:2539-2545(1993).  
RN [17]  
RP VARIANT THR-698.  
RX MEDLINE=91045136; PubMed=2235526;  
RA Zafarullah K., Kleinert C., Tromp G., Kuivaniemi H., Kontusaari S.,  
RA Wu Y., Ganguly A., Prockop D.J.;  
RT "G to A polymorphism in exon 31 of the COL3A1 gene.";  
RL Nucleic Acids Res. 18:6180-6180(1990).  
RN [18]  
RP VARIANT AORTIC ANEURYSM ARG-786.  
RX MEDLINE=91056145; PubMed=2243125;  
RA Kontusaari S., Tromp G., Kuivaniemi H., Romanic A.M., Prockop D.J.;  
RT "A mutation in the gene for type III procollagen (COL3A1) in a family  
RT with aortic aneurysms.";  
RL J. Clin. Invest. 86:1465-1473(1990).  
RN [19]  
RP VARIANT EDS-IV ARG-828.  
RX MEDLINE=94016385; PubMed=8411057;  
RA Richards A.J., Narcisi P., Lloyd J.C., Ferguson C., Pope F.M.;  
RT "The substitution of glycine 661 by arginine in type III collagen  
RT produces mutant molecules with different thermal stabilities and  
RT causes Ehlers-Danlos syndrome type IV.";  
RL J. Med. Genet. 30:690-693(1993).  
RN [20]  
RP VARIANT EDS-IV SER-957.  
RX MEDLINE=89109135; PubMed=2492273;  
RA Tromp G., Kuivaniemi H., Shikata H., Prockop D.J.;  
RT "A single base mutation that substitutes serine for glycine 790 of  
RT the alpha 1 (III) chain of type III procollagen exposes an arginine  
RT and causes Ehlers-Danlos syndrome IV.";  
RL J. Biol. Chem. 264:1349-1352(1989).  
RN [21]  
RP VARIANT EDS-IV VAL-960.  
RX MEDLINE=95268429; PubMed=7749417;

RA Tromp G., de Paepe A., Nuytinck L., Madhathari S.L., Kuivaniemi H.;  
RT "Substitution of valine for glycine 793 in type III procollagen in  
RT Ehlers-Danlos syndrome type IV.";  
RL Hum. Mutat. 5:179-181(1995).  
RN [22]  
RP VARIANT EDS-IV GLU-1014.  
RX MEDLINE=92316511; PubMed=1352273;  
RA Richards A.J., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C.,  
RA Pope F.M.;  
RT "A single base mutation in the gene for type III collagen (COL3A1)  
RT converts glycine 847 to glutamic acid in a family with Ehlers-Danlos  
RT syndrome type IV. An unaffected family member is mosaic for the  
RT mutation.";  
RL Hum. Genet. 89:414-418(1992).  
RN [23]  
RP VARIANT EDS-IV ASP-1050.  
RX MEDLINE=90037070; PubMed=2808425;  
RA Tromp G., Kuivaniemi H., Stolle C.A., Pope F.M., Prockop D.J.;  
RT "Single base mutation in the type III procollagen gene that converts  
RT the codon for glycine 883 to aspartate in a mild variant of  
RT Ehlers-Danlos syndrome IV.";  
RL J. Biol. Chem. 264:19313-19317(1989).  
RN [24]  
RP VARIANT EDS-IV VAL-1077.  
RX MEDLINE=91374480; PubMed=1895316;  
RA Richards A.J., Lloyd J.C., Ward P.N., de Paepe A., Narcisi P.,  
RA Pope F.M.;  
RT "Characterisation of a glycine to valine substitution at amino acid  
RT position 910 of the triple helical region of type III collagen in a  
RT patient with Ehlers-Danlos syndrome type IV.";  
RL J. Med. Genet. 28:458-463(1991).  
RN [25]  
RP VARIANT EDS-IV GLU-1173.  
RX MEDLINE=93022543; PubMed=1357232;  
RA Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.;  
RT "Query Match 9.3%; Score 99; DB 1; Length 1466;  
RT Best Local Similarity 27.8%; Pred. No. 6.5;  
RT Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;  
QY 5 PSDVSSECDREGRLQPAERPPQLRPGAPTSLQT--EP--QGNP-----EGNHGEGG 52  
DB 973 PGQVKGESGKPGANGSLGSRGPPGQGLPLAGTAGEPGRDGNPGDGLPGRDGSPPGKG 1032  
QY 53 DSCPHGSGPGPLAP--PASPGPPATRSPLFIWMRRSSLLSRSSSGYFSFDTDRSPAPMSC 110  
DB 1033 DRGENSGPAGPAGHPGPPGV-----GPGAGKSG 1062  
QY 111 DXSTQTPSPPCQA 123  
DB 1063 DRGESGPAGPAGA 1075  
RESULT 11  
COLL\_HSVS7  
ID COLL\_HSVS7 STANDARD; PRT; 105 AA.  
AC P25050;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Collagen-like protein.  
OS Herpesvirus saimiri (strain 484-77).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Rhadinovirus.  
OX NCBI\_TaxID=10382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90279084; PubMed=2161952;  
RA Geck P., Whitaker S.A., Medveczky M.M., Medveczky P.G.;  
RT "Expression of collagenlike sequences by a tumor virus, herpesvirus  
RT saimiri.";  
RL J. Virol. 64:3509-3515(1990).  
RN [2]

```

RP REVISIONS.
RX MEDLINE=92046383; PubMed=1658399;
RA Geck P., Whitaker S.A., Medveczky M.M., Medveczky P.G.;
RT "Expression of collagenlike sequences by a tumor virus, herpesvirus
RT saimiri";
RL J. Virol. 65:7084-7084(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31964; AAA46152.1; -
DR PIR; A36770; CGBEHS.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF01391; Collagen; 1.
KW Collagen; Repeat.
FT DOMAIN 15 70 COLLAGEN-LIKE.
FT SEQUENCE 105 AA; 10260 MW; EF3DF0FE0FB446F0 CRC64;
CC -----
Query Match 9.2%; Score 98; DB 1; Length 105;
Best Local Similarity 37.2%; Pred. No. 0.58;
Matches 29; Conservative 5; Mismatches 30; Indels 14; Gaps 4;
QY 1 MAKOPS--DVSSECDREGLOQAERPQLRPGAPTSLOTPEQG--MPEGNHGEGDSCP 56
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MASEPNRYRFEETGDRGPGGPGPGPGPG-----PQGGPGPGGPGPGPGPG 52
QY 57 HG--SPQGLAPPASPG 72
Db | |||| | |||
53 QGPPGPGPGPGPGPGPG 70
Db
RESULT 12
CUTL1_HUMAN
ID CUTL1_HUMAN STANDARD; PRT; 1505 AA.
AC P39880; O9UEV5;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CCAAT displacement protein (CDP) (Cut-like 1).
GN CUTL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical vein;
RX MEDLINE=93250973; PubMed=1301999;
RA Neufeld E.J., Skalik D.G., Lievens P.M.-J., Orkin S.H.;
RT "Human CCAAT displacement protein is homologous to the Drosophila
RT homeoprotein, cut.";
RL Nat. Genet. 1:50-55(1992).
RN [2]
RP SEQUENCE OF 48-224 FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
CC -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
CC REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
CC PREVENTING BINDING OF POSITIVELY-ACTIVATING CCAAT FACTORS TO
CC PROMOTERS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;

```

```

CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- DOMAIN: ASN AT POSITION 47 OF THE HOMEBOX MAY PARTICIPATE IN
CC REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
CC HETERODIMERIZATION.
CC -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74099; -: NOT_ANNOTATED_CDS.
DR EMBL; AF047825; AAC78778.1; -
DR HSSP; P10037; IAU7.
DR TRANSFAC; T00100; -.
DR MIM; 116896; -.
DR InterPro; IPR003350; CUT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF02376; CUT; 3.
DR Pfam; PF00046; homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Transcription regulation; Homeobox; DNA-binding;
KW Developmental protein; Nuclear protein; Repeat; Repressor;
KW Coiled coil; Alternative splicing.
FT DOMAIN 7 363 COILED COIL (POTENTIAL).
FT DNA_BIND 542 629 CUT 1.
FT DOMAIN 669 725 COILED COIL (POTENTIAL).
FT DNA_BIND 934 1021 CUT 2.
FT DNA_BIND 1117 1204 CUT 3.
FT DNA_BIND 1244 1303 HOMEBOX.
FT VARSPLIC 632 653 MISSING (IN ISOFORM 2).
FT SEQUENCE 1505 AA; 164353 MW; 860E14D508D4DE11 CRC64;
CC -----
Query Match 9.2%; Score 97.5; DB 1; Length 1505;
Best Local Similarity 23.2%; Pred. No. 8.6;
Matches 52; Conservative 25; Mismatches 82; Indels 65; Gaps 10;
QY 6 SDVSECDREGLOQAERPQLRPGAPTSLOTPEQGNHGGEGDSCPHGSPQ---- 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 SDSLGSARRKGKQ-QPESRRPGSLPAPPSPQLPRNPGQASNTNGTHQFSPAGLSQDFFS 462
QY 62 ----GPIAPPASPGPFATRSPLFIFMRR--SSLLSR-----SSGYFSFD--TD 102
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 SSLASPSLPLASTGKFAINS---LLQRQLMQSFYSKAMQAEAGTSMIFSTGYSTNSS 519
QY 103 RSPAPMSCDKSTQTPSP-----COAFNHVLS 129
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 QSPQQSPDVNGMAPSPSQSESAGSVSEGEDMTAEATARQVKQELIKHNIGQRIFGHTVL 579
QY 130 AMASMRQAEADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQ 173
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
580 GLSGSVSEIL-ARPKPNKLTVR--GKE--PFHKMKQFLSDEQ 618
CC -----
RESULT 13
DLG2_RAT
ID DLG2_RAT STANDARD; PRT; 852 AA.
AC Q63622; Q62939; P70548;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Channel associated protein of synapse-110 (Chapsyn-110) (Synaptic
DE density protein PSD-93) (Discs, large homolog 2).
GN DLG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96310881; PubMed=8755482;  
RA Kim E., Cho K.O., Rothschild A., Sheng M.;  
RT "Heteromultimerization and NMDA receptor-clustering activity of  
Chapsyn-110, a member of the PSD-95 family of proteins.";  
RL Neuron 17:103-113(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96193770; PubMed=8625413;  
RA Brenman J.E., Chao D.S., Gee S.H., McGee A.W., Craven S.E.,  
RA Santillano D.R., Wu Z., Huang F., Xia H., Peters M.F.,  
RA Froehner S.C., Bredt D.S.;  
RT "Interaction of nitric oxide synthase with the postsynaptic density  
protein PSD-95 and alpha1-syntrophin mediated by PDZ domains.";  
RL Cell 84:757-767(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Irie M., Hata Y., Takai Y.;  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR  
SUBUNITS AS WELL AS POTASSIUM CHANNELS.  
CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U49049; AAB53243.1; .  
DR EMBL; U50717; AAC52643.1; .  
DR EMBL; U53368; AAB48562.1; .  
DR HSSP; Q12959; 1PDR.  
DR InterPro; IPR000619; Guanylate\_kin.  
DR InterPro; IPR001478; PDZ.  
DR InterPro; IPR001452; SH3.  
DR Pfam; PF00625; Guanylate\_kin; 1.  
DR Pfam; PF00595; PDZ; 3.  
DR Pfam; PF00018; SH3; 1.  
DR SMART; SM00072; Gukc; 1.  
DR SMART; SM00228; PDZ; 3.  
DR SMART; SM00326; SH3; 1.  
DR PROSITE; PS00856; GUANYLATE\_KINASE\_1; 1.  
DR PROSITE; PS50052; GUANYLATE\_KINASE\_2; 1.  
DR PROSITE; PS50106; PDZ; 3.  
DR PROSITE; PS50002; SH3; 1.  
DR SH3 domain; Repeat.  
KW PDZ 1.  
FT DOMAIN 98 184  
FT DOMAIN 193 279  
FT DOMAIN 421 501  
FT DOMAIN 536 606  
FT DOMAIN 662 852  
FT DOMAIN 181 182  
FT CONFLICT 128 228  
FT CONFLICT 326 326  
FT CONFLICT 339 339  
FT CONFLICT 450 454  
FT CONFLICT 464 465  
FT CONFLICT 474 474  
FT CONFLICT 476 476  
FT CONFLICT 478 478  
FT CONFLICT 484 486  
FT CONFLICT 506 506  
FT CONFLICT 569 569  
FT CONFLICT 586 586  
FT PDZ 2.  
FT PDZ 3.  
FT SH3.  
FT VR -> IL (IN REF. 2).  
FT I -> M (IN REF. 2).  
FT R -> K (IN REF. 2).  
FT D -> E (IN REF. 3).  
FT MISSING (IN REF. 2).  
FT GD -> RK (IN REF. 2).  
FT D -> H (IN REF. 2).  
FT R -> P (IN REF. 2).  
FT A -> D (IN REF. 2).  
FT AAA -> LP (IN REF. 2).  
FT A -> S (IN REF. 2).  
FT H -> N (IN REF. 2).  
FT L -> Q (IN REF. 2).

FT CONFLICT 626 641 MISSING (IN REF. 2).  
FT CONFLICT 639 726 F -> A (IN REF. 3).  
FT CONFLICT 726 733 F -> L (IN REF. 1).  
FT CONFLICT 733 749 N -> Y (IN REF. 2).  
FT CONFLICT 749 756 E -> V (IN REF. 1).  
FT CONFLICT 756 792 L -> H (IN REF. 2).  
FT CONFLICT 791 792 KR -> NG (IN REF. 2).  
FT CONFLICT 794 794 T -> M (IN REF. 2).  
SQ SEQUENCE 852 AA; 94934 MW; F8D414A8B9CF5B09 CRC64;  
  
Query Match 9.1%; Score 97; DB 1; Length 852;  
Best Local Similarity 23.8%; Pred. No. 5.3;  
Matches 44; Conservative 21; Mismatches 58; Indels 62; Gaps 9;  
  
QY 28 LRPGAPTSI-QTEPCQNPENHGEGDSCPH--GSPQGPL-----APPASPGPFA----- 74  
DB 274 LKVGKPTTIYTDTPYGPDPDTHSYSPPMENHLLSGNNGTLEYKTSLPPIISGRYSIPKX 333  
  
QY 75 -----TRSPLFIFMRRSSLLSRSSSGYVFSFDTDRSPAPMSCKS--TOTSPPPQOAFN 125  
DB 334 MLVEDDYTRPPPEPVYSTVKNKCDKPKAS-----PRHYSPEVCDKSFLLSTPY----- 380  
  
QY 126 HYL-----SAMASMRQAPADMRPEIWIQAELRRIGDEFNAYYARVFLNNYQAAEDHP 179  
DB 381 HVHGLLPDSDMTSHSOHSTATRQPSVTLQRAI-----SLEGEPE 419  
  
QY 180 RVVIL 184  
DB 420 RKVVL 424  
  
RESULT 14  
DLG2\_HUMAN STANDARD; PRT; 870 AA.  
AC Q15700;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Channel associated protein of synapse-110 (Chapsyn-110) (Discs, large  
DE homolog 2).  
GN DLG2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96310881; PubMed=8755482;  
RA Kim E., Cho K.O., Rothschild A., Sheng M.;  
RT "Heteromultimerization and NMDA receptor-clustering activity of  
Chapsyn-110, a member of the PSD-95 family of proteins.";  
RL Neuron 17:103-113(1996).  
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR  
SUBUNITS AS WELL AS POTASSIUM CHANNELS.  
CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U32376; AAB04949.1; .  
DR HSSP; Q12959; 1PDR.  
DR MIM; 603583; .  
DR InterPro; IPR000619; Guanylate\_kin.  
DR InterPro; IPR001478; PDZ.



```
DR InterPro; IPR001452; SH3.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF0018; SH3; 1.
DR SMART; SM00072; GuKc; 1.
DR SMART; SM00228; PDZ; 3.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS00502; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
DR SH3 domain; Repeat.
KW DOMAIN 98 PDZ 1.
FT DOMAIN 193 279 PDZ 2.
FT DOMAIN 421 501 PDZ 3.
FT DOMAIN 536 606 SH3.
FT DOMAIN 680 870 GUANYLATE_KINASE.
SQ SEQUENCE 870 AA; 97500 MW; 89CB3BA0619F6F59 CRC64;

Query Match 9.0%; Score 96; DB 1; Length 870;
Best Local Similarity 23.0%; Pred.No. 6.4;
Matches 42; Conservative 24; Mismatches 59; Indels 58; Gaps 8;

QY 28 LRPQAPSL-QTEPQNGNEHGEGDCPH--GSPQGL-----APPASGPFA----- 74
D 1 : | | | : | | | : | | | | | | | | | | | | | | | | | | | |
Db 274 LKVNPTTIYMTDPYGPDIHYSPPMENHLLSGNNGTLEYKTSLPPISPGRYSPIPKH 333
QY 75 -----TRSPLETFEMRRSSLLSRSSCYFSFDTRSPAPMSCDKSTQTPSPPCQANHY 127
D 1 : | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 334 MLVDDDYTRPEPVYSIVNKLCDKPAS-----PRHYSPEVCDKSFLLSAP-----YSHY 382
QY 128 L-----SAMASQROABPADMRPIWIAQELRRIGDEFNAYRVRFLNNYQAAEDHPRM 181
D 1 : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 HGLLPDSEMTSHQHSATAPQSPKTLQRAV-----SLEGEPRK 421
QY 182 VIL 184
D 1 : |
Db 422 VVL 424

RESULT 15
ID CAIH_MOUSE STANDARD; PRT; 1527 AA.
AC P39061; Q62002; Q61437;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].
GN COL18A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=94245707; PubMed=8188673;
RA Rehn M.V., Hintikka E., Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen,
RT partial structure of the corresponding gene, and comparison of the
RT alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen
RT chain.";
RL J. Biol. Chem. 269:13929-13935(1994).
RN [2]
RP SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA Rehn M., Hintikka E., Pihlajaniemi T.;
RT "Characterization of the mouse gene for the alpha-1 chain of type
RT XVIII collagen (COL18A1) reveals that the three variant N-terminal
RT polypeptide forms are transcribed from two widely separated
RT promoters.";
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
```

SEQUENCE OF 213-1140 FROM N.A. (SHORT ISOFORM).  
MEDLINE=94240112; PubMed=8183894;  
Rehn M.V., Pihlajaniemi T.;  
"Alpha 1(XVIII), a collagen chain with frequent interruptions in the  
collagenous sequence, a distinct tissue distribution, and homology  
with type XV collagen.";  
Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).  
[4]  
SEQUENCE OF 240-1527 FROM N.A.  
TISSUE=Liver;  
MEDLINE=94240111; PubMed=8183893;  
Oh S.P., Kamagata Y., Muragaki Y., Timmons S., Ooshima A., Olsen B.R.;  
"Isolation and sequencing of cDNAs for proteins with multiple domains  
of Gly-Xaa-Yaa repeats identify a distinct family of collagenous  
proteins.";  
Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233(1994).  
[5]  
CHARACTERIZATION OF ENDOSTATIN, AND PARTIAL SEQUENCE.  
MEDLINE=97160848; PubMed=9008168;  
O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S.,  
Flynn E., Birkhead J.R., Olsen B.R., Folkman J.;  
"Endostatin: an endogenous inhibitor of angiogenesis and tumor  
growth.";  
Cell 88:277-285(1997).  
[6]  
X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.  
MEDLINE=98169382; PubMed=9501087;  
Hohenester E., Sasaki T., Olsen B.R., Timpl R.;  
"Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A  
resolution.";  
EMBO J. 17:1656-1664(1998).  
CC - FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL  
CC PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY  
CC BINDING TO THE HEPARAN SULPHATE PROTEOGLYCAN INVOLVED IN GROWTH  
CC FACTOR SIGNALLING.  
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM AND A LONG FORM  
CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC - PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING  
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.  
CC - SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH  
CC INTERRUPTED HELICES (FACIT) FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; L16896; AAA37434.1; -  
DR EMBL; U03714; AAA20657.1; -  
DR EMBL; U03715; AAC52901.1; -  
DR EMBL; U34606; AAC52901.1; JOINED.  
DR EMBL; U34608; AAC52901.1; JOINED.  
DR EMBL; U34609; AAC52901.1; JOINED.  
DR EMBL; U34610; AAC52901.1; JOINED.  
DR EMBL; U34611; AAC52901.1; JOINED.  
DR EMBL; U34612; AAC52901.1; JOINED.  
DR EMBL; U34613; AAC52901.1; JOINED.  
DR EMBL; U03716; AAC52901.1; JOINED.  
DR EMBL; U03718; AAC52901.1; JOINED.  
DR EMBL; U03715; AAC52902.1; -  
DR EMBL; U34607; AAC52902.1; JOINED.  
DR EMBL; U34608; AAC52902.1; JOINED.  
DR EMBL; U34609; AAC52902.1; JOINED.  
DR EMBL; U34610; AAC52902.1; JOINED.  
DR EMBL; U34611; AAC52902.1; JOINED.  
DR EMBL; U34612; AAC52902.1; JOINED.  
DR EMBL; U34613; AAC52902.1; JOINED.  
DR EMBL; U03716; AAC52902.1; JOINED.  
DR EMBL; U03718; AAC52902.1; JOINED.  
DR EMBL; U11636; AAC52178.1; -



```
DR EMBL; L22545; AA19787.1; -
DR PDB; 1KOE; 16-FEB-99.
DR MGD; MGI:88451; Col18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF02210; TSPN; 2.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
KW 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 1527
FT CHAIN 1344 1527
FT CHAIN 27 538
FT DOMAIN 539 565
FT DOMAIN 566 575
FT DOMAIN 576 649
FT DOMAIN 650 673
FT DOMAIN 674 795
FT DOMAIN 796 818
FT DOMAIN 819 901
FT DOMAIN 902 915
FT DOMAIN 916 957
FT DOMAIN 958 970
FT DOMAIN 971 1043
FT DOMAIN 1044 1053
FT DOMAIN 1054 1086
FT DOMAIN 1087 1098
FT DOMAIN 1099 1122
FT DOMAIN 1123 1129
FT DOMAIN 1130 1181
FT DOMAIN 1182 1194
FT DOMAIN 1195 1212
FT DOMAIN 1213 1527
FT CARBOHYD 338 338
FT CARBOHYD 700 700
FT DISULFID 1376 1516
FT DISULFID 1478 1508
FT SITE 1104 1106
FT VARSPLIC 1 212
FT VARSPLIC 213 238
FT CONFLICT 900 900
FT CONFLICT 947 947
FT CONFLICT 964 964
FT CONFLICT 1157 1157
FT CONFLICT 1266 1266
FT CONFLICT 1276 1276
FT CONFLICT 1437 1437
SQ SEQUENCE 1527 AA; 156008 MW; 9645045AF140B513 CRC64;

POTENTIAL.
COLLAGEN ALPHA 1(XVIII) CHAIN.
ENDOSTATIN.
NONHELICAL REGION 1 (NC1).
TRIPLE-HELICAL REGION 1 (COL1).
NONHELICAL REGION 2 (NC2).
TRIPLE-HELICAL REGION 2 (COL2).
NONHELICAL REGION 3 (NC3).
TRIPLE-HELICAL REGION 3 (COL3).
NONHELICAL REGION 4 (NC4).
TRIPLE-HELICAL REGION 4 (COL4).
NONHELICAL REGION 5 (NC5).
TRIPLE-HELICAL REGION 5 (COL5).
NONHELICAL REGION 6 (NC6).
TRIPLE-HELICAL REGION 6 (COL6).
NONHELICAL REGION 7 (NC7).
TRIPLE-HELICAL REGION 7 (COL7).
NONHELICAL REGION 8 (NC8).
TRIPLE-HELICAL REGION 8 (COL8).
NONHELICAL REGION 9 (NC9).
TRIPLE-HELICAL REGION 9 (COL9).
NONHELICAL REGION 10 (NC10).
TRIPLE-HELICAL REGION 10 (COL10).
NONHELICAL REGION 11 (NC11).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).

CELL ATTACHMENT SITE (POTENTIAL).
MISSING (IN SHORT ISOFORM).
AVPTQLPPFQSNLQAPLGRPSAPPDF -> MAPRWHLLDVL
TSLVLLLVARVSWAE (IN SHORT ISOFORM).
P -> L (IN REF. 4).
P -> F (IN REF. 4).
A -> R (IN REF. 4).
R -> P (IN REF. 4).
P -> L (IN REF. 4).
L -> F (IN REF. 4).
L -> V (IN REF. 4).

Query Match 9.0%; Score 96; DB 1; Length 1527;
Best Local Similarity 31.9%; Pred. No. 11;
Matches 36; Conservative 5; Mismatches 40; Indels 32; Gaps 6;

QY 16 GROLQPAERPPQLRPGAPTSLQTEPGNPEGNHGGSDSCPFG-----SPQGLAPPA 68
|| : | | | | | | | | | | | | | | | | | | | |
DB 594 GRDGEFGDGEDGRPG-----DTGPGQFP---GTPGDVGPKEGKDPGIGRGPDPGP 644

QY 69 SPGPFATRSPL-FIFMRRSLLSRSSGYFSFDTDRSPAPMSCDKSTQTPSP 120
||| : | | | | | | | | | | | | | | | | | | | |
DB 645 PPGPSRQDKLTFIDMEGSG-----FSGDIESLRGRG-----FPGPP 682
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 10:01:02 ; Search time 26.66 Seconds  
(without alignments)  
1284.809 Million cell updates/sec

Title: US-09-508-832-10  
Perfect score: 1065  
Sequence: 1 MAKOPSDVSSCEDREGRLQ.....PRMVLRLRLRVIVRLVWRMH 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_prodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	110.5	10.4	472	12	O36424	O36424 alcelaphine
2	106.5	10.0	620	12	P89435	P89435 herpes simp
3	106.5	10.0	783	4	Q9NXXD	Q9NXXD homo sapien
4	106.5	10.0	950	5	Q9VE45	Q9VE45 drosophila
5	105.5	9.9	574	3	Q36027	Q36027 schizosacch
6	103.5	9.7	556	2	Q92FS8	Q92FS8 streptomyc
7	103.5	9.7	556	2	Q9S2C0	Q9S2C0 streptomyc
8	103.5	9.7	1051	4	O9S135	O9S135 homo sapien
9	102.5	9.6	1517	11	Q91ZD2	Q91ZD2 mus musculu
10	102	9.6	321	4	Q96SQ7	Q96SQ7 homo sapien
11	101	9.5	1820	13	Q91907	Q91907 pagrus majo
12	100.5	9.4	620	12	Q89893	Q89893 human herpe
13	99.5	9.3	280	5	Q9NFZ9	Q9NFZ9 ascaris suu
14	99.5	9.3	378	11	Q9CSY5	Q9CSY5 mus musculu
15	99.5	9.3	874	3	Q96VJ2	Q96VJ2 pneumocysti
16	99	9.3	257	4	Q9UDU7	Q9UDU7 homo sapien

Q96ha7 homo sapien  
Q94gk7 oryza sativ  
O62432 caenorhabdi  
Q60735 mus musculu  
Q95043 homo sapien  
Q93174 arabidopsis  
Q93107 acanthamoeb  
Q9q49 mus musculu  
Q94dl2 oryza sativ  
Q61434 mus musculu  
Q62001 mus musculu  
Q00487 hydra sp. m  
Q97737 bos taurus  
Q9f691 streptococ  
Q9m33 mus musculu  
Q60749 mus musculu  
O88539 mus musculu  
Q923d5 mus musculu  
Q9jka7 rattus norv  
Q98tw1 brachydanio  
Q9erc1 rattus norv  
O22514 santalum al  
Q9vpp1 drosophila  
Q96c67 homo sapien  
Q99495 homo sapien  
Q99621 homo sapien  
Q84187 measles vir  
Q99in4 canis famil

ALIGNMENTS

RESULT 1  
O36424 PRELIMINARY; PRT; 472 AA.  
ID O36424  
AC O36424;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)  
DE PUTATIVE GLYCOPROTEIN.  
OS Alcelaphine herpesvirus 1 (wildebeest herpesvirus).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Rhadinovirus.  
OX NCBI\_TaxID=35252;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C500;  
RX MEDLINE=97404659; PubMed=9261371;  
RA Esser A., Pflanz R., Fleckenstein B.;  
RT "Primary structure of the alcelaphine herpesvirus 1 genome.";  
RL J. Virol. 71:6517-6525(1997).  
DR EMBL; AF005370; AAC58121.1; -.  
SQ SEQUENCE 472 AA; 50142 MW; 0DFF854764B8342 CRC64;

Query Match 10.4%; Score 110.5; DB 12; Length 472;  
Best Local Similarity 22.9%; Pred. No. 0.0067;  
Matches 48; Conservative 19; Mismatches 78; Indels 65; Gaps 7;  
QY 25 PPOLRPGAPTSL-----QTEPGNPEGNH----- 48  
Db 279 PGCTKPCPTPLWDPCAVLLSSGSTPPWDRRKLRHTKVKTTGTPTVTPVYLISDDDL 338  
QY 49 ---GGEGSDSCPHGSPGGLAPPASPGFATRSPLFIFMRSSLLSRSSSGYFSDTDRSP 105  
Db 339 DRKGAAG-----GGPKKPLPSRDPAGSGQGRG-----TPQSKESTSSGLSGATPRSK 386  
QY 106 APMSCKDKSTPTSPPCQAFNHLYSAMASMRQAPDMRPEIW-IAQLRRIGDFENYAYA 164  
Db 387 APAGRGQRQEPPLPP-----RDYDPASPEALRLRLDIVQKRVYIGGAYNELEA 435



RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cusack S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003723; AAF5584.1; -  
DR FlyBase: FBgn038642; CG7709.  
DR InterPro: IPR002965; P\_rich\_extensin.  
DR InterPro: IPR004019; YLP.  
DR Pfam: PF02757; YLP; 3.  
DR PRINTS: PR01217; PRICEXTENSIN.  
SQ SEQUENCE 950 AA; 91280 MW; B57AB0EA4D30D3D7 CRC64;

Query Match 10.0%; Score 106.5; DB 5; Length 950;  
Best Local Similarity 30.3%; Pred. No. 0.038;  
Matches 37; Conservative 14; Mismatches 58; Indels 13; Gaps 4;  
QY 30 PGAPTSLOTEPGNPNBGGEDSCPHGSPQGLAPPSP--GPATPSPLFIEMRRSS 87  
Db 785 PPAPESSYAAP--SPSSNGGYPAPAPSNSYAGTAPPSSSYCAPSGSPSPSSSY 842  
QY 88 LLRSRSSGGYFSDTRSP-APMSCKSTQTPPPCQAF-----NHVLSAMASMRQAE 138  
Db 843 GAPSTCGSGSSSSSSSSSFGASSSSSAGYPSAPSSSYCAPSTGSGHSFSSAPSSSYAP 902  
QY 139 PA 140  
Db 903 PA 904  
RESULT 5  
ID O36027 PRELIMINARY; PRT; 574 AA.  
AC O36027;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DE WISKOTT-ALDRICH SYNDROME PROTEIN HOMOLOG 1.

GN WSP1 OR SPAC4F10.15C.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=JS21;  
RA Zankel T.C., Ow D.W.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RC SEQUENCE FROM N.A.  
RP STRAIN=972;  
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: TO YEAST LAS17.  
DR EMBL: AF038575; AAB92587.1; -.  
DR EMBL: Z98980; CAB11718.1; -.  
DR InterPro: IPR000697; RanBPL\_WASP.  
DR InterPro: IPR001960; WH1.  
DR InterPro: IPR003124; WH2.  
DR Pfam: PF00568; WH1; 1.  
DR Pfam: PF02205; WH2; 1.  
DR SMART: SM00461; WH1; 1.  
DR DOMAIN 311 317 POLY-PRO.  
FT DOMAIN 337 343 POLY-PRO.  
FT DOMAIN 361 366 POLY-PRO.  
FT VARIANT 248 248 L -> V (IN STRAIN JS21).  
SQ SEQUENCE 574 AA; 59605 MW; C6E5EFC6A02F0E4 CRC64;  
Query Match 9.9%; Score 105.5; DB 3; Length 574;  
Best Local Similarity 26.8%; Pred. No. 0.027;  
Matches 41; Conservative 17; Mismatches 66; Indels 29; Gaps 6;  
QY 21 PAERPPQLRGAPTSLOTEPGQNPGE-----NHGEGDSCPHGSPQGLAPPSPGP--- 72  
Db 424 PSAPPSLPSPAPPSL---PMGAPAAPPLPPSAPIAPPLPAGMPAAPPLPAAAPPPAP 480  
QY 73 -PATRSPPLIFM-----RRSLLSRSSGGYFSDTRSPAPMSCKSTQTPSPCQA- 123  
Db 481 APAPAAPVASIAELPOODGRANLMASIRASGGMDLKRKVSASPSVASTKTNPPVEAP 540  
QY 124 -FNHLSAMAS-----MRQAPADMRIEIV 147  
Db 541 PSNLMALASALNQKTKVAQSDDEDDDEW 573  
RESULT 6  
ID Q92FS8 PRELIMINARY; PRT; 556 AA.  
AC Q92FS8;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE SERINE/THREONINE PROTEIN KINASE.  
GN PKSC.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA Bakal C.J., Davies J.E.;  
RT "Cloning, nucleotide sequence and expression of a serine/threonine  
protein kinase gene from *Streptomyces coelicolor*.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL: AF094711; AAC64406.1; -.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.

QY 1 MAKOPSD-----VSSECDREGQLQPAAERPPOLRPGCAPTSLSQTPEQGPNEN----HG 4  
 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 268 LKKNPNERFSAEAMRDCELRVAASFQA--PSPIVPGAQTS-----SGAGVGSVAFFPV 320

---

QY 50 GEGDSCPHGSPGGPLAPPASGPGR-----ATRSPLFIEMRRSSLLSRSSSGVFSDTD 102  
 :||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 321 GGTPPATPGVPOTPYQPTPSGGPNPYGTTPAASHPAIGYPQQ-----AGT---QT 368

---

QY 103 RSPAPMSCDKSTOTPSPPCQAFNHYLSAMAS 133  
 :|||: |:|:|:|:|:|:|:~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 369 PAPAPYAQQQAATAPPP-----YNLTPSAQCS 395

---

```

RESULT      8
O95135      PRELIMINARY;          PRT;   1051 AA.
AC O95135;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ATAXIN-2-LIKE PROTEIN A2LP.
GN A2LG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CN NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Xia J.-h., Liu C.-y., Wang D.-a., Ruan Q.-g., Deng H.-x.;
RT "A A splicing form of human ataxin-2 like gene obtained from adult
RL brain.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF034373; AAC69607.1; ~~~~~
DR InterPro: IPR002965; Prich_extensn.
DR PRINTS: PR01217; PRICHEXTENSX.
SQ SEQUENCE 1051 AA; 111048 MW; EBB21D51A17FD6B9 CRC64;
```

---

```

Query Match               9.7%; Score 103.5; DB 4; Length 1051;
Best Local Similarity    32.0%; Pred. No. 0.086;
Matches 40; Conservative 8; Mismatches 40; Indels 37; Gaps
```

---

QY 14 REGLOQPAERPPOLRPGCAPTS----LOTPEQNPEGNHGEGDSCPHGSPGGPLAPPAS 69  
 Db 7 RDAPHFPFAARPPPLPRGASSRGFLSPPP-----TRGSPRPPTAGPGC 52  
 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:

---

QY 70 PGPFATRS-----PLTFIMRRSSLRSRSSG-----YFSFTDRSPAPMSCDKST-Q 115  
 :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
 Db 53 PIPLASRALQGPSGPL-----SPSLSNSPSGRCHVEASAATTALPAPAAMAHPATCR 107

---

QY 116 TPSP 120  
 :|:|  
 Db 108 GPSAP 112

---

```

RESULT      9
O91ZD2      PRELIMINARY;          PRT;   1517 AA.
AC O91ZD2;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE CCAT DISPLACEMENT PROTEIN CDF.
GN CUTLI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21429086; PubMed=11544187;
RA Ellis T., Gambardella L., Horcher M., Tschanz S., Capol J.,
```

Qy 161 -----AYARRVFLNNYQAAEDHPRMVILRL-LRYIVRL 193  
|: | | : | : : | | : | | : |

OC	Betaherpesvirinae; Roseolovirus.
OX	NCBI_TaxID=10368;
RN	[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=95266321; PubMed=7747482;  
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,  
RA Martin M.E., Efstathiou S., Craxton M., Macaulay H.A.;  
RT "The DNA sequence of human herpesvirus-6: structure, coding content,  
RT and genome evolution.";  
RL Virology 209:29-51(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=90080132; PubMed=2152817;  
RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W.,  
RA Barrell B.G.;  
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";  
RL J. Virol. 64:287-299(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=91237802; PubMed=1851860;  
RA Chang C.K., Balachandran N.;  
RA "Identification, characterization, and sequence analysis of a cDNA  
RT encoding a phosphoprotein of human herpesvirus 6.";  
RL J. Virol. 65:2884-2894(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=91333007; PubMed=1651403;  
RA Teo I.A., Griffin B.E., Jones M.D.;  
RA "Characterization of the DNA polymerase gene of human herpesvirus 6.";  
RL J. Virol. 65:4670-4680(1991).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=91226542; PubMed=1851252;  
RA Thomson B.J., Efstathiou S., Honess R.W.;  
RA "Acquisition of the human adeno-associated virus type-2 rep gene by  
RT human herpesvirus type-6.";  
RL Nature 351:78-80(1991).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=91374590; PubMed=1654446;  
RA Martin M.E., Nicholas J., Thomson B.J., Newman C., Honess R.W.;  
RA "Identification of a transactivating function mapping to the putative  
RT immediate-early locus of human herpesvirus 6.";  
RL J. Virol. 65:5381-5390(1991).  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=92333249; PubMed=1321206;  
RA Efstathiou S., Lawrence G.L., Brown C.M., Barrell B.G.;  
RA "Identification of homologues to the human cytomegalovirus US22 gene  
RT family in human herpesvirus 6.";  
RL J. Gen. Virol. 73:1661-1671(1992).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=92148942; PubMed=1310766;  
RA Geng Y., Chandran B., Josephs S.F., Wood C.;  
RA "Identification and characterization of a human herpesvirus 6 gene  
RT segment that trans activates the human immunodeficiency virus type 1  
RT promoter.";  
RL J. Virol. 66:1564-1570(1992).  
RN [9]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=93091236; PubMed=1333836;  
RA Gompels U.A., Carss A.L., Sun N., Arrand J.R.;  
RA "Infectivity determinants encoded in a conserved gene block of human  
RT herpesvirus-6.";  
RL DNA Seq. 3:25-39(1992).  
RN [10]

RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=92260671; PubMed=1374813;  
RA Neipel F., Ellinger K., Fleckenstein B.;  
RA "Gene for the major antigenic structural protein (p100) of human  
RT herpesvirus 6.";  
RL J. Virol. 66:3918-3924(1992).  
RN [11]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=92333248; PubMed=1321205;  
RA Thomson B.J., Honess R.W.;  
RA "The right end of the unique region of the genome of human herpesvirus  
RT 6 U1102 contains a candidate immediate early gene enhancer and a  
RT homologue of the human cytomegalovirus US22 gene family.";  
RL J. Gen. Virol. 73:1649-1660(1992).  
RN [12]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=93187613; PubMed=8383182;  
RA Ellinger K., Neipel F., Foa-Tomasi L., Campadelli-Fiume G.,  
RA Fleckenstein B.;  
RT "The glycoprotein B homologue of human herpesvirus 6.";  
RL J. Gen. Virol. 74:495-500(1993).  
RN [13]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=93224882; PubMed=8385692;  
RA Gompels U.A., Carrigan D.R., Carss A.L., Arno J.;  
RA "Two groups of human herpesvirus 6 identified by sequence analyses of  
RT laboratory strains and variants from Hodgkin's lymphoma and bone  
RT marrow transplant patients.";  
RL J. Gen. Virol. 74:613-622(1993).  
RN [14]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=93389439; PubMed=8397282;  
RA Liu D.X., Gompels U.A., Nicholas J., Lelliott C.;  
RA "Identification and expression of the human herpesvirus 6 glycoprotein  
RT H and interaction with an accessory 40K glycoprotein.";  
RL J. Gen. Virol. 74:1847-1857(1993).  
RN [15]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=94025558; PubMed=7692666;  
RA Liu D.X., Gompels U.A., Foa-Tomasi L., Campadelli-Fiume G.;  
RA "Human herpesvirus-6 glycoprotein H and L homologs are components of  
RT the gp100 complex and the gH external domain is the target for  
RT neutralizing monoclonal antibodies.";  
RL Virology 197:12-22(1993).  
RN [16]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=93331710; PubMed=7687803;  
RA Pellett P., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,  
RA Greenamoyer C., Dambaugh T.R.;  
RA "A strong immunoreactive virion protein of human herpesvirus 6  
RT variant B strain 429: identification and characterization of the gene  
RT and mapping of a variant-specific monoclonal antibody reactive  
RT epitope.";  
RL Virology 195:521-531(1993).  
RN [17]  
RP SEQUENCE FROM N.A.  
RC STRAIN=U1102;  
RX MEDLINE=93323202; PubMed=7687301;  
RA Pfeiffer B., Berneman Z.N., Neipel F., Chang C.K., Tirwatnapong S.,  
RA Chandran B.;  
RT "Identification and mapping of the gene encoding the glycoprotein  
RT complex gp82-gp105 of human herpesvirus 6 and mapping of the  
RT neutralizing epitope recognized by monoclonal antibodies.";  
RL J. Virol. 67:4611-4620(1993).  
RN [18]  
RP SEQUENCE FROM N.A.



```

RC STRAIN=U1102;
RX MEDLINE=95146989; PubMed=7844567;
RA Gompels U.A., Macaulay H.A.;
RT "Characterization of human telomeric repeat sequences from human
RT herpesvirus 6 and relationship to replication.";
RL J. Gen. Virol. 76:451-458(1995).
RN [19]
RP SEQUENCE FROM N.A.
RX STRAIN=U1102;
RC MEDLINE=94047392; PubMed=8230490;
RA Dewhurst S., Dollard S.C., Pellett P.E., Dambaugh T.R.;
RT "Identification of a lytic-phase origin of DNA replication in human
RT herpesvirus 6B strain 229.1";
RL J. Virol. 67:7680-7683(1993).
RN [20]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RA Nicholas J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [21]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=94118404; PubMed=8289364;
RA Nicholas J., Martin M.;
RT "Nucleotide sequence analysis of a 39.5-kilobase-pair region of the
RT genome of human herpesvirus 6 encoding human cytomegalovirus
RT immediate-early gene homologs and transactivating functions.";
RL J. Virol. 68:597-610(1994).
RN [22]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=94202284; PubMed=8151768;
RA Schiwe U., Neipel F., Schreiner D., Fleckenstein B.;
RT "Structure and transcription of an immediate-early region in the human
RT herpesvirus 6 genome.";
RL J. Virol. 68:2978-2985(1994).
RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=94181269; PubMed=8134119;
RA Thompson J., Choudhury S., Kashanchi F., Doniger J., Berneman Z.,
RA Frenkel N., Rosenthal L.J.;

Query Match 9.4%; Score 100.5; DB 12; Length 620;
Best Local Similarity 31.0%; Pred. No. 0.093;
Matches 36; Conservative 10; Mismatches 51; Indels 19; Gaps

Qy 21 PAERPOLRP-----GAPTSLQTEPQGNPENHG-----GEGDSCPHGSPQGPLA 65
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 PAEPTTCRPPRPSSDTPLSAVSRPSAPPVPPPTARVFRFLSSSSSSPSYSP-APLS 544
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 66 PPASPGPEATRPSLPIFMRRSSILS--RSSGY-PSFOTDRSPAPMSCDKSTQPS 118
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 PPSVPSPSPSPFPPIPIKSPGLAKPRVSSGHPAAPPAPSSAPARSERVTSVPS 600
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
Q9NEF9 PRELIMINARY; PRT; 280 AA.
ID Q9NEF9
AC Q9NEF9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE CUTICULAR COLLAGEN (FRAGMENT).
GN COL-D.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RA Di Mito C., Betschart B.;
RT "Identification of two collagen genes from the cuticle of Ascaris

```

```

sum.".
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ249373; CAB85466.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR002486; Col_cuticle_N.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF01844; Col_cuticle_N; 1.
KW Collagen.
FT NON_TER
SQ SEQUENCE      280 AA; 27796 MW; E61086529408BC10 CRC64;

Query Match      9.3%; Score 99.5; DB 5; Length 280;
Best Local Similarity 30.4%; Pred. No. 0.046;
Matches 34; Conservative 9; Mismatches 36; Indels 33; Gaps

QY 15 EGROLOAERPPQLRCAPTSLQTEPOGNP--EGNHGGEDGDCPHGSP--QGFLAPPASP 70
Db 193 EAGOLIFGERPPGPPGPPG--RGPPGPRGPPKGDGNDGPPNGPGRGPPGQN 250
QY 71 GPFATRSPLFIEMRRSLLSSSGYVFSFTDTRSPAPMCDKSTQTPSPQC 122
Db 251 G-----QQGRPG-----EPGQAGAGSGCDH-----CPAR 275

RESULT 14
Q9CSY5 PRELIMINARY; PRT; 378 AA.
ID Q9CSY5
AC Q9CSY5; 2001 (TEmBLrel. 17, Created)
DT 01-JUN-2001 (TEmBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEmBLrel. 17, Last annotation update)
DE 2610034E13RIK PROTEIN (FRAGMENT).
GN 2610034E13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner N., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nitaide I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK011662; BAB27764.1; -.
DR MGD; MGI:1922863; 2610034E13RIK.
FT NON_TER
SQ SEQUENCE      378 AA; 40207 MW; DDD783BA958F104C CRC64;

Query Match      9.3%; Score 99.5; DB 11; Length 378;
Best Local Similarity 31.8%; Pred. No. 0.066;
Matches 42; Conservative 10; Mismatches 31; Indels 49; Gaps

QY 16 GROLQAERPPQ--LRPCAPTSLSQTEPOG-----NPGNHGGEDGDCPHGSPG 62

```

Query Match 9.3%; Score 99.5; DB 11; Length 378;  
Best Local Similarity 31.8%; Pred. No. 0.066;  
Matches 42; Conservative 10; Mismatches 31; Indels 49; Gaps 8;

QY 16 GRQLQPAERPPQ--LRPGAPTSLOTEPQG-----NPEGNHGGEGDSCPHGSPQG 62

Search completed: August 16, 2002, 10:04:56  
Job time: 234 sec